Repeat after us: Syntactic alignment is not partner-specific

Rachel Ostrand[a],[*], Victor S. Ferreira[b]

[a] *Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, #0515, La Jolla, CA 92093-0515, United States*
[b] *Department of Psychology, University of California, San Diego, 9500 Gilman Drive, #0109, La Jolla, CA 92093-0109, United States*

ABSTRACT

Conversational partners match each other's speech, a process known as *alignment*. Such alignment can be *partner-specific*, when speakers match particular partners' production distributions, or *partner-independent*, when speakers match aggregated linguistic statistics across their input. However, partner-specificity has only been assessed in situations where it had clear communicative utility, and non-alignment might cause communicative difficulty. Here, we investigate whether speakers align partner-specifically even without a communicative need, and thus whether the mechanism driving alignment is sensitive to communicative and social factors of the linguistic context. In five experiments, participants interacted with two experimenters, each with unique and systematic syntactic preferences (e.g., Experimenter A only produced double object datives and Experimenter B only produced prepositional datives). Across multiple exposure conditions, participants engaged in partner-independent but not partner-specific alignment. Thus, when partner-specificity does not add communicative utility, speakers align to aggregate, partner-independent statistical distributions, supporting a communicatively-modulated mechanism underlying alignment.

Introduction

Human speech is highly variable. Different speakers, and different tokens produced by the same speaker, vary widely in pitch (e.g., Atkinson, 1976; Hanson & Chuang, 1999; Johnson, 2006), phonology and accent (e.g., Terrell, 1976), speech rate (e.g., Miller, Grosjean, & Lomanto, 1984), and word choice (e.g., Lawless, 1984). One way that listeners can cope with this variability in their input is by encoding and adapting to their interlocutors' idiosyncratic linguistic characteristics (Kleinschmidt & Jaeger, 2015). In turn, speakers frequently modulate their own speech to match characteristics of the linguistic input that they have received, in a process known as *linguistic alignment*. There is substantial evidence that alignment occurs, to some degree, at multiple linguistic levels during every conversation. However, prior research has not addressed the degree to which speakers align to the statistical distribution of the overall, aggregated recent linguistic context, versus to a statistical distribution for each particular interlocutor a speaker interacts with. The present work addresses this question in a series of experiments investigating syntactic alignment between a speaker and two listeners who have distinct production statistics.

The ubiquity and utility of linguistic alignment

Linguistic alignment occurs at nearly every level of language processing, from high-level semantic and discourse processes down to low-level acoustic characteristics and temporal and structural properties of the conversation. For example, alignment has been abundantly demonstrated at the lexical level (often called *lexical entrainment*), as speakers reuse the words produced by their partners when referring to the same object (e.g., Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986; Garrod & Anderson, 1987). Alignment has additionally been demonstrated at the phonetic level, as speakers change their pronunciation to sound more like their partners or model talkers (e.g., Goldinger, 1998; Pardo, 2006, 2010); and at the acoustic level, as speakers align their pitch, acoustic spectra, fundamental frequencies, and loudness to each other (e.g., Gregory, Dagan, & Webster, 1997; Natale, 1975; Shockley, Sabadini, & Fowler, 2004). Speakers also align to temporal characteristics, by converging on their partners' speech rate (e.g., Street, 1984; Webb, 1969), turn duration (e.g., Matarazzo, Weitman, Saslow, & Wiens, 1963), and pause length (e.g, ten Bosch, Oostdijk, & de Ruiter, 2004), and to their partners' gestures (e.g., Bergmann & Kopp, 2012; Holler & Wilkin, 2011; Kimbara, 2008).

[*] Corresponding author at: Computational Biology Center, IBM Research, 1101 Kitchawan Road, Yorktown Heights, NY 10598, United States.
E-mail address: ostrand.rachel@gmail.com (R. Ostrand).

Importantly for present purposes, there is ample evidence that speakers align their syntactic production based on their recent syntactic experience. When speaking with a live interlocutor, speakers are biased toward producing the same structure that their partner just produced (Branigan, Pickering, & Cleland, 2000; Cleland & Pickering, 2003; Haywood, Pickering, & Branigan, 2005; Levelt & Kelter, 1982). Speakers syntactically align not just in experimental tasks but also in unscripted everyday conversation (Gries, 2005; Reitter & Moore, 2014). Speakers are additionally influenced by the overall statistics of their syntactic input, and produce more of a structure that occurred disproportionately often during their recent experience (Kaschak, 2007; Kaschak, Kutta, & Jones, 2011; Kaschak, Loney, & Borreggine, 2006).

The Rational Speaker Model of syntactic alignment

The process of alignment begins when the language processing system adapts its representations of linguistic knowledge to match the statistics of its current (but constantly changing) environment. When the language system receives a particular linguistic feature in its input (e.g., a particular syntactic structure), that indicates the increased prevalence of that feature in the linguistic environment, and so the system comes to expect that feature more in the future. Then, when the speaker produces speech, the system is more likely to produce that feature back, in comparison to if the system had not previously processed that linguistic feature in its input. In short, as the comprehension system learns the statistics of its current environment and adapts its representations to those statistics, the production system reflects that contextual statistical learning in its own subsequent produced utterances. This results in the behavioral effect of alignment.

One account of the syntactic alignment just described (and the more general effect that includes syntactic alignment, namely, *syntactic priming*; see Pickering & Ferreira, 2008, for review) is termed the *Rational Speaker Model* (Fine, Jaeger, Farmer, & Qian, 2013; Jaeger & Snider, 2013). According to the Rational Speaker Model, speakers' representations of linguistic knowledge adapt to the statistics of the current environment for a reason: By adapting to the current contextual statistics, the system can better cope with the wide variability it is faced with when trying to interpret linguistic signals. The syntactic system, for example, is (at least in principle) generative, and so can generate an open-ended and therefore potentially infinite variety of syntactic sequences. By encoding that variability, and attributing it to sources that reliably account for that variability, the system can more efficiently process the range of linguistic signals that it must comprehend (see also Kleinschmidt & Jaeger, 2015).

One very important information source to which variability can be attributed is the individual speaker who generated a linguistic utterance. People speak differently from one another, but a particular speaker should tend to speak in an internally-consistent way. For example, different speakers use different syntactic structures at different frequencies, but a particular speaker likely will use a more consistent or systematic set of syntactic structures. In other words, between-speaker variability is likely larger than within-speaker variability. According to the Rational Speaker Model, the system should adapt to these speaker-specific statistics, by learning a speaker's linguistic distribution and attributing systematic behavior to that speaker, and thus the process of adapting to input statistics is *partner-specific*. A rational speaker should represent that different conversational partners speak in systematically different ways.

At the same time as a rational speaker adapts partner-specifically, they also adapt *partner-independently* (Kleinschmidt & Jaeger, 2015). That is, adaptation is hierarchical, such that the rational speaker represents both the statistics of their individual partners, and the statistics of the general speaker – namely, aggregated across the overall linguistic environment. This model of the general speaker can then be brought to bear on the process of communicating with a brand new speaker for whom there is not yet a partner-specific model (among possible

functions for this general model).

The theoretical claim of the Rational Speaker Model that the process of syntactic adaptation, in the service of syntactic comprehension, has a partner-specific aspect, raises an interesting possibility: Should the process of syntactic *alignment* in production also be partner-specific? That is, when a rational speaker speaks back to a partner whose individualized linguistic distribution they have (hypothetically) learned, does their syntactic production display partner-specificity as well? The answer to this question hinges on how the syntactic production system that enacts syntactic alignment operates. Specifically, upon addressing a partner for whom the system has learned a speaker-specific distribution, if the syntactic production system encodes that it is formulating an utterance for that (adapted-to) partner, then the system could engage in one of two behaviors. The system could use the partner-specific model as its probability distribution for selecting syntactic features for production, which would be manifest behaviorally as partner-specific alignment – that is, the speaker should match each partner's individualized (syntactic) distribution when talking to that partner. Or the system could use the partner-general model as its probability distribution for selecting syntactic features for production, which would be manifest behaviorally as partner-independent alignment – that is, the speaker should speak (syntactically) the same way to all partners, regardless of those partners' individual systematic syntactic behaviors. Alternatively, it is possible that a speaker-specific distribution was never learned in the first place, whether due to insufficient experience, not enough between-speaker variability, within-speaker lack of consistency, or any other reason. Relatedly, the production system might not encode that this utterance is directed to an adapted-to partner. In those situations, the system could only rely on the partner-general model (i.e., the only model that exists) as its probability distribution for selecting syntactic features for production, which would be manifest as partner-independent alignment. However, when both models have been learned, the reason why a rational speaker model might choose the partner-specific versus partner-general model during production is discussed next.

Alignment as a function of communicative utility

In the literature to date, the clearest evidence of partner-specific alignment occurs in the referential domain, in a process called *lexical entrainment*. After a speaker and listener jointly establish shared labels for objects in the current referential context (e.g., "the ice skater" for an abstract geometric design), the speaker continues to use the same labels as long as she is interacting with the same listener. In contrast, when interacting with a new listener who was not present for the previous conversation, speakers revert to a longer and more elaborated reference (e.g., "it looks like a person ice skating who's sticking out their foot"; Brennan & Clark, 1996; Wilkes-Gibbs & Clark, 1992). In fact, speakers avoid the shared but less-informative label ("the ice skater") if there is *any* listener present who was not part of the original conversation and thus does not know the shared label. Speakers do so even if the new partner had heard the original conversation as a bystander (but not an active participant; Wilkes-Gibbs & Clark, 1992), or when addressing the old partner and a new partner together, thus breaking alignment with a knowledgeable partner in order to accommodate a co-present unknowledgeable partner (Yoon & Brown-Schmidt, 2014). Furthermore, after agreeing upon a particular label to refer to an object, listeners expect the original speaker will continue to use that shared label, as listeners take longer to identify an object if an old speaker uses a new expression than if a new speaker uses the identical new expression (Brennan & Hanna, 2009; Metzing & Brennan, 2003).

However, a common thread among these lexical entrainment studies that show partner-specific alignment is *communicative utility* – where a lack of alignment has a good chance (or at least the potential) of causing communicative difficulty. If a speaker says "the ice skater" when describing an abstract shape to a partner who was not present for

the original labelling of the picture, then the listener will have substantial difficulty understanding which picture is being referenced and may fail to do so. Lexical entrainment occurs because conversational partners have created a *conceptual pact* (Brennan & Clark, 1996), in which partners build a shared agreement of how to conceptualize a referent for the purpose of the conversation. When one partner violates that pact by producing a different label, comprehension is disrupted. But if a new partner arrives, she is not a member of the previous conceptual pact, so the original speaker and the new partner must create a new conceptual pact between themselves and establish their own shared lexical label. Additionally, in lexical entrainment experiments, speakers typically describe unusual objects that do not have common names – either a novel object that has no predetermined label in the language, or a low-frequency, subordinate-level object that needs to be differentiated from other objects in the same basic-level category (e.g., "loafer" and "docksider", both subordinate-level descriptions of shoes). Thus, participants in these experiments could assume that a new partner who was not present for the initial labeling of the abstract picture as "the ice skater" might not understand the referent, or in the case of a subordinate label like "the docksider," have more trouble identifying the object, due to its low frequency. As a result, lexical entrainment may be different from alignment at other linguistic levels in that partner-specific behavior reduces communicative difficulty.

In contrast, there are other linguistic features that do not rely on forming a conceptual pact – such as those conveyed by variation in syntactic structure, low-level acoustic properties, and paralinguistic measures – that exhibit context-wide alignment but likely do not raise issues of communicative difficulty. It seems unlikely that producing a syntactic structure or pause duration or fundamental frequency to a new partner would give that partner trouble in understanding the utterance (at least in the partner's native language). On the other hand, speakers do show alignment to the *overall* linguistic context for these features.

The lexical entrainment literature demonstrates that speakers learn and align to partner-specific linguistic distributions when such partner-specificity *is* important for facilitating communication. However, at the other end of the continuum of communicative utility, it is unknown whether alignment is partner-specific or partner-independent. Understanding whether speakers track and align to the linguistic statistics for each individual partner, versus aggregated, partner-independent statistics, when there is less of a communicative need to build a partner-specific shared conceptualization (as is likely the case for the majority of linguistic features) will help to elucidate the mechanism behind linguistic alignment. However, most prior studies investigating alignment have perfectly confounded partner-specific and partner-independent behavior, as their participants only interacted with one partner and that partner either did not have a consistent linguistic behavior (e.g., Branigan et al., 2000; Cleland & Pickering, 2003; Haywood et al., 2005; Holler & Wilkin, 2011; Kaschak, 2007; Kimbara, 2008; Levelt & Kelter, 1982; Matarazzo et al., 1963; Natale, 1975; Pardo, 2006, 2010; Street, 1984; ten Bosch et al., 2004; Webb, 1969) and/or the participant immediately repeated their partner's production, without any intervening context (e.g., Babel, 2012; Goldinger, 1998; Namy, Nygaard, & Sauerteig, 2002; Shockley et al., 2004), in which case "alignment" was measured as the degree to which a participant matched their partner's relevant linguistic feature on the immediately-preceding interaction. Although such behavior demonstrates speakers' ability to learn and adapt to their input, there is no way to determine whether the speaker reproduced a given feature in their own speech because they were talking to the same person who originally produced that feature, or because that feature was present in their immediately-preceding linguistic context. That is, with only one partner, there is no way to disentangle whether participants learn and use in their production a separate distribution for each partner, or an aggregated distribution for the entire linguistic context – because there is only one distribution to be learned, which matches both the current partner *and*

the current context.

That said, there is a small group of studies that deconfounded partner-specific and partner-independent alignment by having participants interact with two partners, and measuring whether participants learned and/or aligned to partner-specific distributions. Among these, results are consistent with the hypothesis that language users track individual partner's production distributions when it could be communicatively useful, although they almost exclusively consist of positive evidence.

When studies have measured linguistic features that could have communicative utility, they have shown partner-specific alignment (in production) or adaptation (in comprehension). As noted above, lexical entrainment work has demonstrated that participants track separate lexical distributions for each partner, expect different words from different partners for the same object (e.g. Brennan & Hanna, 2009; Metzing & Brennan, 2003), and align to their current partner's production (e.g. Brennan & Clark, 1996; Wilkes-Gibbs & Clark, 1992; Yoon & Brown-Schmidt, 2014). These studies used abstract pictures or subordinate-level labels that could be difficult to understand without having been part of the original conversation and thus partner-independent alignment could potentially induce some communicative difficulty.

Additionally, there is evidence of partner-specific phonetic adaptation in comprehension, in which listeners expect different words when they hear the same phonemes from different speakers, based on earlier experience with that speaker (Kraljic & Samuel, 2007; Trude & Brown-Schmidt, 2012). A speaker's pronunciation can have an important role for communicative success or efficiency, especially for minimal pairs – learning that when Speaker A says /beIg/, she means "beg" but when Speaker B produces the same set of phonemes, she means "bag" can affect comprehension and reduce the processing cost of surprisal induced by mis-parsing phonemes into the incorrect word.

There is also initial evidence that listeners can learn partner-specific syntactic distributions in the case of ambiguity. Kamide (2012) exposed listeners to three speakers who produced syntactically ambiguous relative clause attachment sentences (e.g., "The uncle of the girl who will ride the motorbike/carousel…"). One speaker always expressed one interpretation (*high-attachment*, meaning the uncle is riding), another speaker always expressed a different interpretation (*low-attachment*, meaning the girl is riding), and the third speaker used the same ambiguous surface structure for the two meanings equally often. After this exposure, listeners made anticipatory eye-movements to the object that matched the speaker's preferred meaning (motorbike or carousel) when they heard one of the consistent speakers' voices, but not for the inconsistent speaker's voice. This suggests that listeners can learn syntactic distributions tied to a particular partner and use that to drive their comprehension expectations in a partner-specific manner. As the experimental sentences were ambiguous, it is possible that such partner-specific learning occurred due to communicative utility, because learning that one speaker intends one parse for an ambiguous utterance, and a different speaker intends a different parse for the same utterance, could reduce comprehension confusion and garden-pathing. Note, however, that a more recent paper failed to replicate this partner-specific learning, and found that listeners were *not* more likely to look at the object matching the speaker's syntactic preference (Liu, Burchill, Tanenhaus, & Jaeger, 2017), so this result should be further investigated.

In contrast, in the few studies which both (a) investigate linguistic features for which partner-specific learning does *not* have a clear potential communicative benefit and (b) are equipped to test for partner-specific alignment, speakers aligned partner-independently. In a relatively unusual study in the lexical entrainment literature (Mol, Bogers, & Bouwens, 2012), participants alternately interacted with two confederates, describing map landmarks to each other. Although the landmarks were easily-identifiable pictures of different building types (unlike the abstract tangram images or low-frequency nouns used in

many lexical entrainment studies), the confederates used "rather specific wording" (p. 758) such as "historical city palace" or "small intimate terrace" to ensure that if the participants repeated the confederate's words, it was due to alignment instead of accidental repetition. Participants showed partner-independent but not partner-specific alignment – although they reproduced the original picture description ("historical city palace") over 75% of the time, they did so regardless of whether they were speaking to the confederate who originally produced that description or to the other confederate. An important difference between the lexical entrainment studies which *did* show partner-specific alignment and this one is communicative utility – a new partner should be able to identify which building is the "historical city palace" among a set of dissimilar buildings, but may not be able to identify which abstract image is "the ice skater". Although this should be interpreted cautiously, the authors *did* find a marginal effect of partner-specific behavior between participants and confederates performing the same task in their second language, which is consistent with a communicative utility mechanism, as lower proficiency may necessitate greater alignment.

At the syntactic level, Branigan, Pickering, McLean, and Cleland (2007) investigated syntactic production using unambiguous dative sentences (in contrast to the ambiguous sentences used by Kamide, 2012). Participants interacted in triads, which included a speaker (who produced the prime syntactic structure), an addressee (who selected the picture described by the speaker), and a side-listener (who observed the conversation but did nothing). The rate at which participants reused the prime sentence's syntactic structure differed as a function of their own conversational role when the prime sentence had been produced (more priming when they had been the addressee compared to the side-listener of the speaker's prime sentence). However, when holding constant the participant's own conversational role, participants were just as likely to repeat the prime sentence's syntactic structure regardless of whether they were speaking to the person who had produced that prime sentence or a side-listener, demonstrating partner-independent priming. However, in this study, the participants' partners produced both structures (PDs and DOs) equally often, and all three speakers were in the same room and interacted jointly. As a result, there was little communicative pressure for participants to speak differently to their two partners, as both partners had overheard and (in some conditions) had successfully produced, both syntactic structures themselves, thus supporting partner-independent alignment in this case where partner-specific alignment is not needed for communicative success. In fact, given that the two partners did not demonstrate unique, consistent syntactic distributions for the participants to (potentially) learn, it was not even possible for participants to align specifically to their listener's syntactic preference, and alignment was measured not as the participant's learning of each partner's aggregate syntactic statistics, but rather whether the participant's sentence matched the immediately preceding syntactic structure.

A comprehension study (Squires, 2014) addressed this issue by presenting listeners with two talkers who each had unique, consistent syntactic distributions and systematically produced sentences containing either a standard or non-standard use of the verb "don't" (e.g., "The **dogs** don't want food." [standard] vs. "The **dog** don't want food." [non-standard]). Listeners learned which speaker had produced "dog" and which had produced "dogs" for the exact sentences they were exposed to, but did not generalize the partner-specific learning to new sentences. These results suggest that listeners did not learn a generalized linguistic distribution for each partner, but rather associated specific sentences' syntax with specific speakers (perhaps even via explicit paired associations). In contrast to the ambiguous sentences used by Kamide (2012), matching a partner's use of "don't" likely does not affect communicative success and thus there should not be partner-specific learning pressure.

In sum, the literature to date suggests that partner-specific alignment is most prevalent in situations when it is either communicatively necessary or the speaker believes it may aid their listener's comprehension. However, this has not been tested in a linguistic context with two conversational partners who produce internally-consistent but differing linguistic distributions. To fully investigate the nature of alignment and what type of distribution speakers draw from during dialogue, a linguistic situation is needed in which variability between speakers is systematic and easily generalizable to new utterances, such that partner-specific alignment is cognitively possible, but does not necessarily affect communicative success.

To test these questions, the present experiments investigate the partner-specific versus partner-independent nature of alignment in the domain of syntax. A speaker's syntactic production can be systematic (i.e., always producing the same structure for a particular event type, such as double object datives for dative events) and as such it is in principle possible for listeners to generalize from exposed instances to novel productions (and generate their own). Critically, reusing a partner's preferred syntactic structure is unlikely to affect communicative success or utility, as native speakers produce both structures of common syntactic alternations such as transitives and datives by the time they are 3–4 years old (Shimpi, Gámez, Huttenlocher, & Vasilyeva, 2007; Snyder & Stromswold, 1997). As a result, fluent, adult speakers should be familiar with all of the standard syntactic alternatives available in their native language, and listeners should have strong prior expectations that, although their partner preferred one syntactic structure, they could nevertheless successfully comprehend both. This allows us to test whether speakers align partner-independently in a linguistic context where partner-specific behavior does not impact communicative success or utility.

From the perspective of the Rational Speaker Model, because syntactic alignment is likely to be communicatively unnecessary, the syntactic production process should consult the partner-general model that the adaptation process has created, and thus syntactic alignment should be partner-independent. The contrasting possibility is that the syntactic alignment of a rational speaker consults the partner-specific model by default, irrespective of communicative necessity. If so, then even in the syntactic domain, where a lack of partner-specific alignment is unlikely to have communicative consequences, alignment should be partner-specific.

The present research tests these possibilities by decoupling the two types of linguistic exposure that could contribute to alignment – exposure from the current partner and exposure from the current context. Participants received equal exposure to two conversational partners with opposing, and systematic, syntactic production statistics, and participants' degree of alignment to each partner was measured. The present experimental design gives participants the opportunity to engage in either partner-specific or partner-independent alignment even when there is no communicative need to do so, by testing whether speakers reflect distinct, partner-specific statistical distributions or the overall contextual statistics in their own syntactic production.

Present research

In five experiments, participants interacted with two conversational partners who each had distinct and systematic syntactic preferences. After being exposed to both partners' speech, participants then spoke to each partner individually. Participants played a picture-matching game with the two partners, in which one person described a series of scenes while the other selected and ordered the specified pictures. (For ease of exposition, participants will be referred to as male and the conversational partners as female.) The two partners exhibited internally-consistent but contrasting syntactic preferences. For example, when describing dative events, one partner *always* produced prepositional datives (PD; e.g., "The girl threw her sandwich to the lion"), and the other partner *always* produced double object datives (DO; e.g., "The girl threw the lion her sandwich"). Participants listened to one partner at a time use her preferred syntactic structure to describe several pictures,

Table 1
Exposure conditions for each experiment.

	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5
Pictures described by the participant without hearing them described by either experimenter		x	x	x	x
Pictures described first by one experimenter and later by the participant	x	x	x		
Pictures described first by one experimenter, then the other experimenter, then the participant			x		
Between-participants conditions where both experimenters produced the same structure				x	x
Experimenter with lower linguistic proficiency					x

and then participants described several pictures to each partner individually. We measured how often participants produced the syntactic structure that the currently-listening partner produced (exclusively) herself. (See Table 1 for an overview of the conditions in each experiment.)

Experiment 1

Experiment 1 tested for partner-specific alignment in syntax, where it is cognitively possible, but should not affect communicative utility. The experiment provided a rich syntactic environment in which participants interacted with two experimenters who exhibited consistent and differing syntactic preferences for three event types. To make learning the association between an experimenter and her syntactic preference as easy as possible, participants were exposed to an experimenter's syntactic preferences for specific pictures, and then later described the same pictures back. Thus, participants had the opportunity to learn that a particular partner preferred a particular structure for a particular picture, and partner-specific alignment could arise either via the episodic traces of each sentence produced by each partner, or by generalizing across the exposure and abstracting a structural preference for each partner. If alignment is driven by a rote statistical mechanism, and an addressee's prior linguistic behavior impacts their partner's production regardless of communicative utility, then participants should be biased to use their listener's preferred syntactic structure, thus demonstrating partner-specific syntactic alignment. In contrast, if alignment is driven by a mechanism which takes communicative utility into account, then participants should produce structures at the same frequencies to their two partners and engage in partner-independent syntactic alignment.

Method

Participants

Sixty-four students at the University of California, San Diego participated in this experiment for course credit. All reported being native, monolingual speakers of English.

Materials

Stimuli consisted of 48 colored drawings of scenes. Each picture could be described using two different syntactic structures, and was designed to elicit either a simple transitive (described using either an active or passive sentence), a locative (described using either "with" or "on"), or a dative (described using either a prepositional dative [PD] or double object [DO] construction). The stimuli were evenly divided between the three alternations. Each picture was printed on a card approximately 4 1/2" tall and 3 2/3" wide. See Table 2 for example

Fig. 1. Sample dative picture from the present experiments.

descriptions for each type of picture, Fig. 1 for a sample picture, and Appendix A for a complete list of stimuli used in the experiment.

Procedure

Participants were told they were playing a conversational picture-matching game with the experimenters. One participant and one experimenter sat across a table from each other, separated by an opaque barrier that was high enough to block the other's table space but low enough to easily see each other's face and upper body. Each partner had a series of pictures; the task throughout the experiment was for one partner to describe the pictures to the other partner, who put his/her own pictures in the same order.

Each participant interacted with two experimenters, one at a time. Both experimenters were female. A round began when the first experimenter entered the participant's room and described six distinct pictures to the participant (two each of transitive, locative, and dative events), while the participant arranged his own cards in the order described (Exposure Phase A). The first experimenter gave the participant two 2-digit multiplication problems to complete and left the room. The purpose of the math problems was to provide a cover task to allow the experimenter to leave; performance on the math problems was not measured. After 30 seconds, the second experimenter entered the participant's room, collected the math problems, and described a new set of six distinct pictures to the participant (Exposure Phase B). The second experimenter then gave the participant a new pair of math problems and left the room. Finally, one experimenter returned, this time as the listener, and laid out the participant's pictures in a pre-determined and pseudo-random order such that two pictures of the same event type would not be described consecutively. The participant described all 12 pictures that he had just heard (both experimenters' full set of six) to the listening experimenter (Test Phase). Thus, for each picture the

Table 2
Example sentences used by the experimenters to describe the stimuli.

Syntactic construction	Syntactic alternation A	Syntactic alternation B
Transitive	A tornado is destroying a barn. (*active*)	A barn is being destroyed by a tornado. (*passive*)
Locative	Santa Claus is stuffing the stockings with fruit. (*with-locative*)	Santa Claus is stuffing fruit into the stockings. (*on-locative*)
Dative	A boy is tossing his dad the ball. (*double object*)	A boy is tossing the ball to his dad. (*prepositional dative*)

participant described, the listening experimenter was either the same or different person as the experimenter who had originally described that picture to the participant. This process comprised a complete round (Experimenter A described six unique pictures, Experimenter B described six unique pictures, then the participant described the same 12 pictures), and occurred for four rounds, each containing different pictures. Each experimenter described a total of 24 distinct pictures, and the participant described all 48 pictures over the course of the experiment.

All factors (including nuisance factors) were fully counterbalanced either (or both) within or between participants. The order of the two experimenters during each exposure phase, and identity of the experimenter during each test phase, was counterbalanced across rounds for a given participant and also across participants. That is, each participant listened to Experimenter A and then Experimenter B during two exposure rounds, and Experimenter B and then Experimenter A during the other two exposure rounds. Similarly, each participant described his pictures to Experimenter A for two test rounds and to Experimenter B for the other two test rounds. Additionally, picture-structure mapping was counterbalanced across participants, such that half of participants heard (e.g.) Fig. 1 described using a DO, and the other half heard it described using a PD. The order in which each experimenter described her pictures, and the order in which participants described their pictures, was also counterbalanced across participants.

One experimenter described each transitive picture using an active sentence, each locative picture using a with-locative sentence, and each dative picture using a double object sentence. The other experimenter produced only passive, on-locative, and prepositional dative sentences. The identity of the experimenter who had each syntactic preference was also counterbalanced between participants, such that half of the participants heard actives, with-locatives, and DOs from "Hannah", and passives, on-locatives, and PDs from "Victoria", and the other half of participants heard the reversed mapping. At the start of the experiment, participants were told there were multiple experimenters running the experiment at the same time, and thus they might encounter a new experimenter later in the experiment, but were never explicitly informed about each experimenter's syntactic preferences, or that the pictures could be described using multiple syntactic structures.

Analysis

For this and all following experiments, syntactic production was analyzed using generalized logit mixed-effects models using the lme4 package (version 1.1.13; Bates, Maechler, Bolker, & Walker, 2015) in R (version 3.4.0; R Core Team, 2017). Models included maximal random effects and a full correlational structure (following Barr, Levy, Scheepers, & Tily, 2013): All within-subject fixed effects varied by subjects and all within-item fixed effects varied by item. When models failed to converge, the same model-reduction strategy was followed for all experiments until convergence was achieved. First, correlations between random effects were removed. If the model still did not converge, the random effect accounting for the least variance was iteratively removed (with the caveat that random intercepts were always maintained).

All independent variables were categorical, and were sum contrast-coded with factor levels coded as -0.5 and $+0.5$. To determine statistical significance for each effect, the full model was compared with a reduced model that had the effect in question removed; the random effects structure for the two compared models was matched in the case of non-convergence of the sub-model. For independent variables with three or more levels, all levels were removed simultaneously during model comparison to test for the presence of that effect. For example, the effect of Event Type consists of three levels – transitives, locatives, and datives. To test for the main effect of Event Type, two levels must be removed from the model: the level of locative compared to transitive, and the level of dative compared to transitive. Thus, the reported main effect of Event Type will have a two degrees of freedom in the χ^2 test.

As effect sizes are difficult to interpret in the log-odds space of GLMMs, for this and all following analyses, subject means are reported in the text and presented in the associated figures. Subjects who did not contribute data to all conditions were excluded for these descriptive purposes; for the different experiments and analyses reported below, this ranged from 0 to 17 excluded subjects. Note, however, that all subjects were retained in the statistical analyses in the GLMMs.

Data for all experiments reported in this paper are available at: https://osf.io/tafnb/.

Results

Participants' descriptions for each picture were transcribed and coded offline (blind to which experimenter had originally described the picture and which was listening) as one of the two intended structures for that picture, or as a different (other) structure. Note that the participant's description was counted as the relevant alternation type even if it did not exactly match the experimenter's initial description. For example, following the experimenter's description of, "A boy is tossing his dad the ball," one participant later described that picture as, "There's a small boy throwing a ball to an older man." As long as the participant's sentence was a reasonably accurate description of the picture (as all productions were) and the correct alternation (here, a dative sentence), it was included.

For all analyses below, production frequency is reported in terms of the "metric structure" for each event type, chosen to be the structures produced by Experimenter B (passives, on-locatives, and PDs). If participants produce fewer metric structures to Experimenter A (who never produced those structures) than to Experimenter B (who exclusively produced those structures), this would demonstrate partner-specific alignment.

Some productions could not be categorized as one of the two intended structures of an alternation and were removed from the analysis. These included descriptions that did not mention all of the necessary noun arguments (two in the transitive scenes, three in each of the locative and dative scenes, e.g., "The boy is throwing the ball"; 3.9%), or were not a complete sentence of the intended alternation (e.g., "The baseball" or "A boy is playing catch with a man"; 11.1%). These exclusions removed 11.5% of the transitive events, 17.6% of the locative events, and 16.0% of the dative events.

The remaining syntactic productions were analyzed with a 2 (Listener: Experimenter A, Experimenter B) × 3 (Event Type: transitive, locative, dative) logit mixed-effects model. Results are shown in Fig. 2.

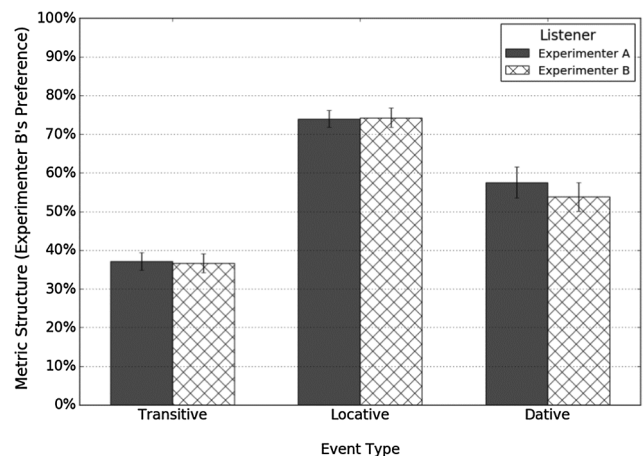Participants did not align their syntactic production to match their

Fig. 2. Results from Experiment 1. Percentage of participants' descriptions that used the metric structure (Experimenter B's preference: passive, on-locative, or PD) for each event type, as a function of the listener's identity. Error bars show standard error of the mean.

listener's preference. They were no more likely to produce, for example, PDs to the PD-preferring experimenter than to the DO-preferring experimenter (metric structures [passives/on-locatives/PDs] to Experimenter A, who never produced these structures: 56.2%; to Experimenter B, who always produced these structures: 54.9%; no main effect of Listener: $\chi^2(1) = 0.798$, $p < .372$). Production varied across the different alternations, such that on-locatives were a larger percentage of the total locative productions (74.1%) compared to PDs of the total dative productions (55.7%) or passives of the total transitive productions (36.9%; main effect of Event Type: $\chi^2(2) = 14.268$, $p < .0008$). However, this effect is of no theoretical interest to the present work, as it indexes the relative bias in English of one structure over its alternate (passives are greatly dispreferred compared to actives; on-locatives are greatly preferred compared to with-locatives; PDs and DOs are relatively balanced) and thus will not be discussed further in this or the following experiments. There was also no interaction ($\chi^2(2) = 0.997$, $p < .607$), meaning that the lack of partner-specific alignment was not modulated by the relative bias of the different alternations.

Across the experiment, no partner-specific alignment was demonstrated. However, participants heard only two sentences of each event type from each experimenter per round. It is possible that participants did not yet have enough experience to learn each experimenter's syntactic preferences in the early part of the experiment. To address this possibility, the omnibus analysis was repeated including only participant productions from the final two rounds. Thus this analysis includes participants' productions after they had heard, respectively, six and eight consistent syntactic structures of each event type from each experimenter, an amount that has been shown to induce partner-specific adaptation at other linguistic levels (e.g., Kraljic & Samuel, 2007; Metzing & Brennan, 2003; Trude & Brown-Schmidt, 2012). The results remained the same: Participants were equally likely to produce metric structures to the experimenter who had never produced those structures (58.1%) as to the experimenter who had always produced those structures (56.5%; no main effect of Listener: $\chi^2(1) = 0.010$, $p < .920$). Contrary to the full data set, there was no main effect of Event Type ($\chi^2(2) = 1.562$, $p < .458$), likely caused by variability in structure bias for different pictures, as a given picture always appeared in the same round across subjects; thus half of the items were excluded in this analysis. There was a marginal interaction between Listener and Event Type ($\chi^2(2) = 5.455$, $p < .065$), which was driven by a larger production of PD sentences when speaking to the DO-preferring experimenter (59.7% PDs) compared production of PD sentences when speaking to the PD-preferring experimenter (49.9% PDs). Given the tenuous nature of this result, the fact that it goes in the opposite direction as would demonstrate partner-specific alignment, and the fact that there was no such difference in production for the other two event types, this marginal effect is likely due to simply a noisy pattern of data engendered by halving the number of observations. Thus overall, even when restricted to the rounds in which participants had received the most exposure to their partner's preferences, there was no effect of partner-specific alignment.

However, it is possible that the lack of observed alignment occurred because participants were not attending to the task or were unable to remember information about the individual pictures. Thus, to verify that participants had paid attention to the experimental context and could modify some aspect of their syntactic production in response to their syntactic input, we turned to recent work (Gruberg, Ostrand, Momma, & Ferreira, 2019), which shows that speakers reuse the same syntactic structure that they heard a particular picture previously described with. For example, if a participant heard the picture shown in Fig. 1 described using a PD, he is more likely to reuse a PD (rather than DO) structure to describe that *specific picture*, even after several minutes and intervening sentences (including other dative sentences). This observation, which we refer to here as *event-specific alignment*, provides a manipulation check in the present experiment of whether participants

paid attention to the task and were able to track some aspects of syntactic variability. (There are several possible mechanisms which could induce event-specific alignment in this experiment, including episodic traces of the original description or its syntactic structure. For present purposes, the exact mechanism driving event-specific alignment is not important; its relevance to the current work is merely as a manipulation check.)

To test for event-specific alignment, participants' descriptions (the same descriptions as were analyzed to test for partner-specific alignment) were categorized based on which structure had been originally used by the experimenter to describe that picture, and submitted to a 2 (Original Structure: non-metric [actives, with-locatives, DOs], metric [passives, on-locatives, PDs]) × 3 (Event Type: transitive, locative, dative) generalized logit mixed-effects model. There was a robust effect of event-specific alignment: For the pictures which an experimenter had previously described with a metric structure, participants were significantly more likely to also describe that particular picture using a metric structure (62.8%) compared to those pictures which an experimenter had previously described with a non-metric structure (47.8%; main effect of Original Structure: $\chi^2(1) = 120.23$, $p < .0001$). There was a main effect of Event Type, indexing the relative biases of different structures used in the experiment ($\chi^2(2) = 14.858$, $p < .0006$). There was also an interaction, as there was a greater effect of event-specific alignment for transitives (28.5% effect) than locatives (13.2% effect) and datives (3.5% effect; $\chi^2(2) = 60.609$; $p < .0001$).

Discussion

Experiment 1 tested whether speakers align their syntactic choices to match their partners' consistent syntactic preferences. There was no evidence of partner-specific alignment. After interacting with two speakers who each exhibited robust syntactic preferences, participants did not tailor their own speech to use the structures that their current listener preferred: They were no more likely to produce a particular structure when it matched their listener's own syntactic production than when it mismatched. These results support a mechanism which modulates alignment as a function of communicative utility, as it is unlikely that participants thought their partners could not understand the referent of a non-aligned syntactic structure.

However, participants did demonstrate event-specific alignment by repeating the syntactic structure that they had previously heard used to describe a given picture. For example, participants who heard the picture shown in Fig. 1 described by Experimenter B using a PD structure were likely to describe that picture back as a PD, regardless of whether they were speaking to Experimenter A (who preferred DOs) or Experimenter B (who preferred PDs). This effect provides evidence that participants tracked some characteristics of their syntactic input and modified their syntactic production in response to speech they heard within the experiment. As a result, participants were more likely than not to repeat the structure they heard as that picture's earlier description, regardless of their listener's demonstrated syntactic preference. In the present experiment, event-specific and partner-specific alignment acted partially in opposition. If a participant heard a picture described by Experimenter A using a DO, when he described it later, he will be biased to use a DO as well. Half of the time, the listener is Experimenter A, and the structural association from the picture works in favor of partner-specific alignment, as the picture primes the listener's preferred structure (DO). But half of the time, the listener is Experimenter B, in which case the structural association from the picture works against partner-specific alignment, as the picture primes the DO but the listener's preferred structure is the PD. As every picture was first described by one of the experimenters and later described back by the participant, the potential influence of event-specific alignment on all of the trials in this experiment could have reduced the observable effects of partner-specific alignment. This concern was addressed in Experiment 2.

Experiment 2

In Experiment 2, participants again interacted with two experimenters who produced differing syntactic structures for each of three event types. This experiment addressed whether speakers align to their partner's syntactic preferences when there is no event-specific influence. To that end, Experiment 2 included two groups of pictures for participants to describe: pictures that were first described by one experimenter, and pictures that were not described by either experimenter. If syntactic alignment is partner-specific, but was blocked by the event-specific alignment in Experiment 1, then participants should use their listeners' preferred structure when describing pictures they have not previously heard, but continue to use the structure from the original description for pictures they were previously exposed to.

Method

Participants

Sixty-four students at the University of California, San Diego who did not take part in Experiment 1 participated in this experiment for course credit. All were native, monolingual speakers of English.

Materials

The stimuli were 72 pictures of the same type as those used in Experiment 1 (including many of the Experiment 1 stimuli). There were 24 transitive, 24 locative, and 24 dative event pictures. See Appendix A for the complete list of stimuli.

Procedure

The procedure was identical to that of Experiment 1 with one important difference: Not all of the pictures were initially described by one of the experimenters. Instead, 1/3 of the pictures were described by the participant but neither experimenter (No Experimenter); 1/3 of the pictures were described first by one experimenter and later the participant (One Experimenter); and, for counterbalancing purposes, 1/3 of the pictures were described by an experimenter but not the participant (No Participant). Half of the picture descriptions to each participant (equally divided across the One Experimenter and No Participant scenes) were described by Experimenter A using an active, on-locative, or DO structure, and the other half were described by Experimenter B using a passive, with-locative, or PD structure. The dependent measure was again the frequency of production of metric structures (Experimenter B's preferences). As in Experiment 1, the experiment consisted of four rounds; during each round, Experimenters A and B each described six pictures, and then the participant described 12 pictures to one listening experimenter. Participants described half of their pictures (equally divided across the No Experimenter and One Experimenter conditions) to Experimenter A and the other half to Experimenter B, with the listener order counterbalanced across participants. Both experimenters were female.

Results

Participants' descriptions were transcribed and coded offline (blind to the previous description exposure condition and listening experimenter's identity) as one of the two intended structures for that picture, or neither. Trials which were not the intended alternation (20.6%), were missing a necessary noun argument (1.4%), or were affected by an experimenter error (0.1%) were excluded. These exclusions removed 17.1% of transitives, 26.7% of locatives, and 22.8% of datives. One item was subsequently excluded because it had data points from fewer than 30% of participants. The remaining responses were submitted to a 2 (Listener: Experimenter A, Experimenter B) × 2 (Previous Description: No Experimenter, One Experimenter) generalized logit mixed-effects model. Participant means are reported below; results are shown in Fig. 3.
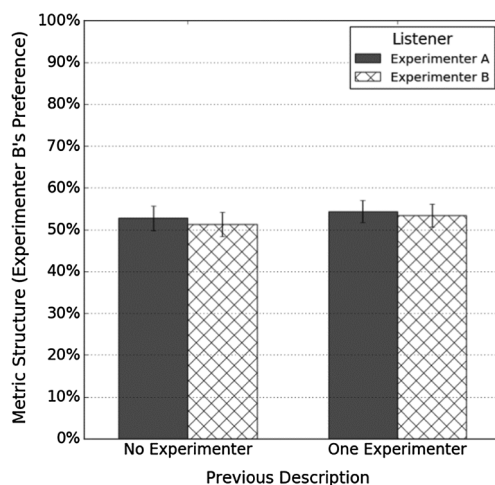

Fig. 3. Results from Experiment 2. Percentage of participants' descriptions that used the metric structure (Experimenter B's preference: passive, on-locative, or PD) for each event type, as a function of whether the pictures had been previously described by one of the experimenters and the listener's identity. Error bars show standard error of the mean.

As in Experiment 1, participants did not show partner-specific alignment. They were no more likely to produce a particular structure to the experimenter who preferred that structure than to the experimenter who dispreferred it (metric structures to Experimenter A, who never produced these structures: 53.5%; to Experimenter B, who always produced these structures: 52.4%; no main effect of Listener: $\chi^2(1) = 0.254$, $p < .615$). Participants produced metric structures overall at the same rate when describing pictures that they had never heard before (No Experimenter: 52.0% metric structures) compared to pictures that one of the experimenters had previously described (One Experimenter: 53.9% metric structures; no main effect of Previous Description: $\chi^2(1) = 0.422$, $p < .516$). Importantly, there was no modulation of partner-specific alignment as a function of whether a picture had been previously described or not. Participants were just as likely to produce a particular structure when talking to the experimenter who preferred it as who dispreferred it both for the No Experimenter pictures (to Experimenter A: 52.7% metric structures; to Experimenter B: 51.3% metric structures) and for the One Experimenter pictures (to Experimenter A: 54.3% metric structures; to Experimenter B: 53.4% metric structures; no Listener × Previous Description interaction: $\chi^2(2) = 0.081$, $p < .776$).

To test for partner-specific alignment after maximal exposure to each partner's syntactic distribution, this analysis was repeated in a 2 (Listener: Experimenter A, Experimenter B) × 2 (Previous Description: No Experimenter, One Experimenter) GLMM restricted to descriptions from the third and fourth rounds of the experiment. (The same one item was excluded as in the full analysis, for having an acceptable description from fewer than 30% of participants.) The results mirrored those of the full analysis: Participants did not speak differently to each partner (to Experimenter A: 52.0% metric structures; to Experimenter B: 52.0% metric structures; no main effect of Listener: $\chi^2(1) = 1.006$, $p < .316$). Participants spoke the same way to their two partners both when describing pictures they had not heard described before (to Experimenter A: 49.8%; to Experimenter B: 46.3%) and pictures they had heard one experimenter describe (to Experimenter A: 54.2%; to Experimenter B: 57.8%; no Listener × Previous Description interaction: $\chi^2(1) = 1.720$, $p < .190$ or main effect of Previous Description: $\chi^2(1) = 0.674$, $p < .412$)

However, participants did show event-specific alignment, and reused the structure that a particular picture was described with earlier in the experiment when they described that picture back. Productions in the One Experimenter condition were submitted to a 2 (Original

Structure: metric, non-metric) × 3 (Event Type: transitive, locative, dative) GLMM. (Because the pictures in the No Experimenter condition had not been previously described by an experimenter, participants could not match their syntactic structure to that picture's original structure.) Participants used a metric structure significantly more often when that picture had previously been described using a metric structure (62.1%) than when that picture had previously been described using a non-metric structure (52.3%; main effect of Original Structure: $\chi^2(1) = 43.306$, $p < .0001$). There was a main effect of Event Type ($\chi^2(2) = 11.075$, $p < .004$), indexing the relative biases of different structures used in the experiment. The degree of event-specific alignment differed between the three Event Types, with more event-specific alignment for locatives (18.6% effect) than transitives (14.7% effect) than datives (-3.9% effect; Original Structure × Event Type interaction: $\chi^2(2) = 7.896$, $p < .019$). This clear effect of event-specific alignment demonstrates that participants were paying attention to the task and the experimenters, and were able to modify their syntactic production based on their syntactic input.

Discussion

Experiment 2 further tested whether speakers' syntactic choices are influenced by the linguistic exposure they receive from their partners. Consistent with the results from Experiment 1, participants were no more likely to produce a particular syntactic structure when speaking to the listener who preferred that structure compared to the listener who dispreferred it. Thus, there was no evidence for partner-specific alignment, supporting the claim that the mechanism driving alignment takes its communicative utility into account. However, participants always heard the two experimenters describe different pictures; perhaps this lack of explicit contrast made participants think the experimenters' syntactic preferences were internally consistent just by coincidence. Experiment 3 addressed this possibility by having both experimenters describe the same pictures with their own preferred syntactic structure.

Experiment 3

In the previous experiments, although participants had the potential to learn that their two interlocutors had differing syntactic preferences, the differences were never directly contrasted, as a given picture was described by at most one experimenter. Perhaps the alignment mechanism assigned the previously-used structure as that picture's optimal description. If this was the case, then the inference may be that Experimenter A happened to describe all of the pictures that were intrinsically DO-biased, and Experimenter B happened to describe all of the pictures that were intrinsically PD-biased. If so, a rational participant should not infer that the experimenters' syntactic consistency signaled something about their own preferences; instead, a rational participant should infer that the experimenters' syntactic regularity was merely coincidental, and did not signal a meaningful linguistic preference. In Experiment 3, the contrast between the experimenters' syntactic preferences was highlighted and this possible inference eliminated.

Participants once more interacted with two experimenters who each had unique and internally-consistent syntactic preferences. In addition to the conditions from the previous two experiments, this experiment added pictures that were described by both experimenters, each using their own preferred structure. In this condition, a participant heard the identical picture described by Experimenter A as, "A boy is tossing his dad the ball" and by Experimenter B as, "A boy is tossing the ball to his dad." This gave the participant explicit evidence of (a) each experimenter's syntactic preference on a particular picture and (b) that a particular picture could be acceptably described in two ways and each experimenter was producing one structure at the expense of another one.

Method

Participants

Sixty-four students at the University of California, San Diego who did not take part in Experiments 1 or 2 participated in this experiment for course credit. All reported being native, monolingual speakers of English.

Materials

The stimuli were 48 pictures of the same type as those used in Experiments 1 and 2, equally split between transitive, locative, and dative events. See Appendix A for the complete list of stimuli.

Procedure

The experimental procedure was identical to that of Experiment 2 with one additional picture condition: pictures that were described first by one experimenter (e.g., "A boy is tossing the ball to his dad.") and then later in the same round, again by the other experimenter (e.g., "A boy is tossing his dad the ball.") Additionally, the No Participant filler condition was removed for this experiment. Therefore, this experiment consisted of three Previous Description conditions: No Experimenter (as in Experiment 2), One Experimenter (as in Experiments 1 and 2), and Two Experimenters (new for this experiment). As in the previous two experiments, each participant heard half of the pictures from Experimenter A (who preferred the non-metric structures: actives, with-locatives, and DOs) and half from Experimenter B (who preferred the metric structures: passives, on-locatives, and PDs), and then described half of his pictures to each experimenter. All factors were counterbalanced either (or both) between or within subjects. The dependent measure was the frequency of production of metric structures. There were three experimenters, all female; different participants were tested by different pairs of experimenters, assigned non-systematically based on availability.

Results

Participants' descriptions were transcribed and coded offline (blind to the previous description exposure condition and listening experimenter's identity) as one of the two intended structures for each alternation, or neither. Trials on which the participant did not produce the intended alternation (11.1%) or was missing one of the necessary noun arguments (2.3%) were excluded. These exclusions removed 13.6% of transitives, 18.8% of locatives, and 7.7% of datives. The remaining productions were submitted to a 2 (Listener: Experimenter A, Experimenter B) × 3 (Previous Description: No Experimenter, One Experimenter, Two Experimenters) GLMM. Participant means are reported; results are shown in Fig. 4.

Even with specific evidence of the experimenters' contrasting syntactic preferences, there was no evidence of partner-specific alignment. Participants did not produce different structures as a function of their listener's identity (metric structures to Experimenter A: 48.0%; to Experimenter B: 45.7%; no main effect of Listener: $\chi^2(1) = 0.047$, $p < .828$). Participants were equally likely to produce metric structures when describing pictures that had been previously described by No Experimenter (46.5%), by One Experimenter (48.1%), or by Two Experimenters using different structures (45.8%; no main effect of Previous Description: $\chi^2(2) = 0.095$, $p < .954$). There was an equivalent lack of partner-specific alignment regardless of how often the participant had heard that picture previously described (no Listener × Previous Description interaction: $\chi^2(2) = 0.449$, $p < .799$).

Participants' productions from only the final two rounds of the experiment were analyzed in a 2 (Listener: Experimenter A, Experimenter B) × 2 (Previous Description: No Experimenter, One Experimenter, Two Experimenters) GLMM to see if alignment behavior emerged after increased exposure. Participants still did not speak differently to the two experimenters (to Experimenter A: 50.9% metric structures; to Experimenter B: 45.7% metric structures; no main effect of Listener: $\chi^2(1) = 2.342$,
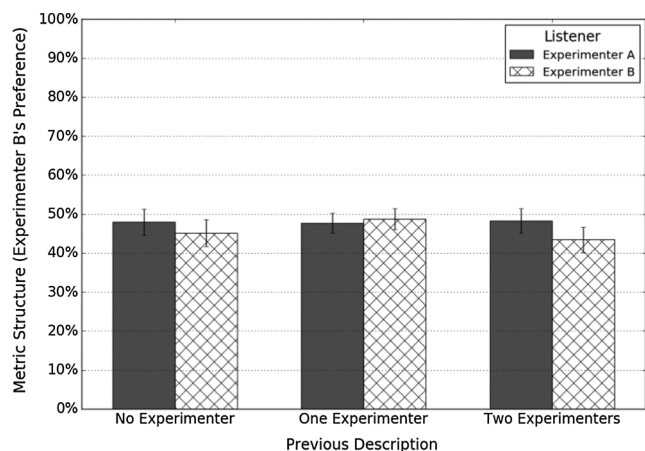
Fig. 4. Results from Experiment 3. Percentage of participants' descriptions that used the metric structure (Experimenter B's preference: passive, on-locative, or PD) for each event type, as a function of whether the pictures were previously described by zero, one, or two experimenters, and the listener's identity. Error bars show standard error of the mean.

$p < .126$). (Note that the slight numerical effect of Listener – which did not even reach marginal significance – is in the direction of anti-alignment: Participants produced slightly numerically more of Experimenter B's preferred structures when speaking to Experimenter A.) There was no difference in alignment behavior based on how participants had interacted with the picture earlier in the experiment (No Experimenter: 48.6%; One Experimenter: 48.8%; Two Experimenters: 47.5%; no main effect of Previous Description: $\chi^2(2) = 3.327$, $p < .190$) and no interaction ($\chi^2(2) = 0.617$, $p < .617$).

To verify that participants modulated their syntactic production based on some aspects of their syntactic input, event-specific alignment was assessed by submitting productions from the One Experimenter condition to a 2 (Original Structure: non-metric, metric) × 3 (Event Type: transitive, locative, dative) GLMM. (Pictures in the No Experimenter and Two Experimenters conditions did not have a unique original structure and thus it is not meaningful to test for the presence of event-specific alignment for these items.) Participants engaged in robust event-specific alignment: They produced more metric structures when that picture was previously described with a metric structure (56.3%) than with a non-metric structure (39.7%; main effect of Original Structure: $\chi^2(1) = 40.146$, $p < .0001$). There was a marginally significant main effect of Event Type ($\chi^2(2) = 5.627$, $p < .060$), and participants showed more event-specific alignment for transitive pictures (22.6% effect) and locative pictures (19.7% effect) than dative pictures (7.4% effect; Original Structure × Event Type interaction: $\chi^2(2) = 7.287$, $p < .026$). These data provide a reliability check that participants were paying attention to their interlocutors and to the task, and could modulate their syntactic production based on syntactic input received within the context of the experiment.

Discussion

The results from Experiment 3 support those of Experiments 1 and 2. When describing scenes, participants did not align their syntactic production to match the structures that the listening experimenter had previously produced. Rather, participants were equally likely to produce a given structure (e.g., a PD sentence) when talking to the experimenter who had produced exclusively PDs for dative events as when talking to the experimenter who had produced exclusively DOs for dative events. This lack of partner-specific syntactic alignment held for pictures for which participants had no event representation (from never hearing them described); for pictures with a single event representation (from hearing them described by one experimenter); and

for pictures with two event representations, providing explicit evidence that they could be described with two structures (from hearing them described by both experimenters using different syntactic structures). These latter pictures also gave the participant experience that each partner produced one structure at the expense of the other.

However, participants repeated the structure for pictures they had heard described exactly once. So whereas there was no evidence for partner-specific syntactic alignment, this experiment once again demonstrated event-specific syntactic alignment and speakers' ability to modulate their syntactic production in response to their input.

The lack of partner-specific syntactic alignment has been replicated across three experiments, including when restricting the analysis to sentences produced after having received most of the exposure to their partners' preferences. However, participants heard only eight sentences of each event type from each experimenter. Although listeners adapt to partner-specific distributions after a similar amount of exposure at other levels of language processing (e.g., Kraljic & Samuel, 2007; Metzing & Brennan, 2003; Trude & Brown-Schmidt, 2012), it is possible that this is not enough exposure to learn or align to each experimenter's syntactic distribution. Experiment 4 assessed whether the lack of partner-specific syntactic alignment that was observed in Experiments 1–3 is the result of insufficient syntactic experience or truly indicative of partner-independent alignment.

Experiment 4

Experiment 4 had two goals: one theoretical and one practical. The theoretical goal was to investigate whether speakers aligned to aggregated, partner-independent statistical distributions. The practical goal was to address the possibility that the previous experiments did not give participants sufficient exposure to their partners' syntactic preferences. To achieve these two goals, we made three major methodological modifications to the experimental design. First, we tripled the number of sentences of a particular structure that participants heard. Second, we tested whether participants could learn and explicitly recall the content of statements spoken by each experimenter. Third, we included control conditions to measure global, partner-independent syntactic alignment (following Kaschak, 2007; Kaschak et al., 2011; Kaschak et al., 2006), testing whether participants could modify their syntactic production after a given number of exposure trials.

Method

Participants

Ninety-six students at the University of California, San Diego who did not take part in any of the previous experiments participated in this experiment for course credit. All reported being native, monolingual speakers of English.

Materials

The stimuli were 96 colored pictures in the same style as those in Experiments 1–3, consisting of 72 unique dative pictures and 24 unique intransitive pictures. As before, the dative events could be described using either a prepositional dative (PD) or double object (DO) structure. For item counterbalancing purposes, the dative pictures were divided into three item sets of 24. For a particular participant, Experimenter A described one set of dative pictures, Experimenter B described a different set, and the participant described the third set.[1] Across participants, the items sets were counterbalanced among the three speakers in

---

[1] As discussed below, participants in the speaker-specific and mixed conditions heard all 24 dative pictures in one set described by Experimenter A, and all 24 dative pictures in another set described by Experimenter B. However, participants in the all-PD and all-DO conditions only heard 12 of the 24 pictures in each set described by, respectively, Experimenters A and B.

Table 3
Between-participant exposure conditions in Experiments 4 and 5. PD = prepositional dative; DO = double object; int = intransitive.

	Syntax exposure group			
	All-PD	Speaker-specific	Mixed	All-DO
Exp'er A says:	12 PD + 12 int	24 DO	12 PD + 12 DO	12 DO + 12 int
Exp'er B says:	12 PD + 12 int	24 PD	12 PD + 12 DO	12 DO + 12 int

the experiment, so that a particular picture was described by Experimenter A for 1/3 of participants, described by Experimenter B for 1/3 of participants, and described by the participant for 1/3 of participants.

In certain conditions, the intransitive pictures were described by the experimenters as filler items, to hold constant the number of exposure sentences of a given dative structure across the conditions (see Table 3, below). Additionally, all participants described the intransitive pictures interleaved with the critical dative pictures, to reduce the influence of self-priming from one sentence to the next. The intransitive pictures had a simple event structure (e.g., "The woman is sleeping") which made it unlikely that participants would produce a sentence containing even a single object, so as not to prime one of the dative structures more than the other. See Appendix A for the full list of stimuli used in the experiment.

Procedure
Main experiment. The general procedure was similar to that of the previous experiments, with a few important differences. As before, each subject interacted with two experimenters across the experiment, and only one experimenter was ever in the room with the subject at a time. Across participants, there were seven people who acted as experimenters, six female and one male; different participants were tested by different pairs of experimenters, assigned non-systematically based on availability.

One goal for Experiment 4 was to ensure that the experimental manipulation was sufficiently powerful to detect partner-specific syntactic alignment should such alignment exist. This was addressed with four methodological changes from the previous experiments. First, to increase the amount of exposure to each experimenter's syntactic preferences, transitive and locative pictures were removed so that all critical trials involved dative pictures, thus increasing the number of sentences of one alternation (PD vs. DO) that participants heard from the experimenters. Second, for the same reason, all of the rounds in which the experimenter described her pictures preceded all of the rounds in which the participant described his pictures. This maximized the amount of syntactic exposure that participants received from each experimenter before describing their own pictures. As a result of these two design changes, participants heard 24 sentences of a given structure (PD or DO) before they described any pictures themselves.

Third, to verify that participants were aware that they were interacting with two distinct experimenters and could remember specific statements that were said by each person, each round began with the experimenter telling the participant a fictional but plausible fact about herself – for example, that the experimenter grew up in New York City (an uncommon occurrence among students attending a public university in California). Participants were told at the beginning of the experiment that the facts would "come up later", and were tested on which experimenter had said each fact after the main experiment.

The fourth departure from previous experiments was that each participant was randomly assigned to one of four between-participant syntax exposure groups. As before, each experimenter described a total of 24 pictures across four rounds to the participant. In the *all-PD* exposure condition, both experimenters described all of their dative events using PDs. Thus participants in this condition heard a total of 24 PDs, 0 DOs, and 24 intransitives (12 PDs and 12 intransitives from each experimenter). The *all-DO* exposure condition was the reverse: both

experimenters described all of their dative events using DOs. Thus, in this condition, participants heard a total of 0 PDs, 24 DOs, and 24 intransitives (12 DOs and 12 intransitives from each experimenter). Comparing the relative rate of participants' PD and DO production in these two conditions will permit testing for global syntactic alignment. If hearing 24 sentences of a given alternation is sufficient to affect participants' syntactic production, then participants in the all-PD condition should produce more PDs than do participants in the all-DO condition (following Kaschak, 2007). In the *mixed* condition, both experimenters produced both structures at the same rate, each describing half of her dative events using PDs and half using DOs. Thus, participants heard a total of 24 PDs, 24 DOs, and 0 intransitives (12 PDs and 12 DOs from each experimenter). In the critical, *speaker-specific* exposure condition, participants heard the same overall number of sentences of each structure as in the mixed condition (24 PDs, 24 DOs, and 0 intransitives). However, each experimenter described all of her pictures using only her preferred structure. Thus, participants in this condition heard all 24 DOs from Experimenter A and all 24 PDs from Experimenter B. Across the four syntax exposure conditions, all participants received the same amount of exposure to a particular structure – 24 sentences. If hearing 24 sentences of one structure is sufficient to affect syntactic production in one condition, it should be for the others as well. See Table 3 for a summary of the exposure conditions.

Exposure phase. Aside from these changes in design, the procedure for Experiment 4 was very similar to that of the previous experiments. The experiment began with a short practice round. Next, one experimenter told the participant a fact about herself and engaged him in a short chat relating to the fact (e.g., "Purple is my least favorite color"), and then described six cards to the participant according to that participant's exposure condition. The experimenter gave the participant a 2-digit multiplication problem and left the room for 15 seconds. The second experimenter then entered the room, told the participant a different fact about herself (e.g., "I grew up in New York City"), described a different set of six cards to the participant, gave him another math problem, and left the room. This process constituted one round (Experimenter A: fact, six picture descriptions, math problem; Experimenter B: fact, six different picture descriptions, math problem). There were four rounds in the exposure phase, and every picture was unique. Thus, participants heard a total of eight facts (four from each experimenter) and 48 picture descriptions (24 from each experimenter). Order of experimenters within each round (Experimenter A vs. Experimenter B first), picture set, and picture order were all counterbalanced between participants. In the speaker-specific and mixed exposure conditions, every picture described by the experimenters was a dative event. In the all-PD and all-DO exposure conditions, half of the pictures described by each experimenter were dative events, and half were intransitive events (in order to give every participant equivalent exposure to one structure). In all four exposure conditions, the first six pictures and the last six pictures in the exposure phase were always dative events, to equate any potential effects of primacy or recency across conditions.

Test phase. Following the four rounds of experimenter descriptions, the participant described 12 pictures (six datives and six intransitives, interleaved) to one of the experimenters. Then the experimenter gave the participant a math problem and left the room. The second

experimenter entered, and the participant described another 12 pictures (again, six datives and six intransitives, interleaved) to her, and completed another math problem. The first experimenter returned and the participant described another 12 pictures and completed a math problem, and then the second experimenter returned and the participant described a final 12 pictures and completed one more math problem. Thus, each participant described 24 scenes (12 datives, 12 intransitives) to each experimenter. All dative pictures that the participant described to the experimenters were unique and had not been described by either experimenter. The intransitive pictures that the participant described were the same intransitive pictures that participants in the all-PD and all-DO exposure conditions had heard from the two experimenters. (Note that participants in the speaker-specific and mixed exposure conditions heard only dative pictures and thus no intransitive pictures from either experimenter.) The order of the listening experimenters and the order of the pictures was counterbalanced between participants.

Post-experiment questionnaire. Following the main experiment, participants completed a questionnaire on the computer, which queried participants' memory for the facts and syntactic structure(s) each experimenter produced. For details on the questionnaire, see the Supplementary Materials.

Results

Main experiment

Participants' descriptions were transcribed and coded offline (blind to syntax exposure condition and listener identity) as one of the two dative alternations (PD or DO), or neither. Descriptions that could not be categorized as one of the alternations, either because they were not a dative structure (26.0%) or were missing an argument (6.8%) were excluded. Seven pictures (out of 72) were excluded from the analyses because they had observations from fewer than 30% of subjects.[2]

The included sentence productions were submitted to a 4 (Syntax Exposure Condition: all-PD, speaker-specific, mixed, all-DO) × 2 (Listener: Experimenter A, Experimenter B) GLMM. Participant means are reported below, and shown in Fig. 5. If participants received enough exposure to influence their partner-independent, global syntactic production, there should be higher PD production in the all-PD condition compared to the all-DO condition. If participants aligned partner-specifically to the experimenters, there should be differential PD production when speaking to the two experimenters in the speaker-specific condition, but equivalent PD production when speaking to the two experimenters in the mixed condition.

There was a significant main effect of Syntax Exposure Condition ($\chi^2(3) = 8.042$, $p < .045$). Participants produced the highest proportion of PDs after hearing only PDs from both experimenters (all-PD exposure condition: 64.5% PDs), an intermediate amount after hearing both PDs and DOs (speaker-specific: 62.2% PDs; mixed: 53.6% PDs), and the lowest proportion of PDs – and thus the highest proportion of DOs – after hearing only DOs from both experimenters (all-DO: 45.2% PDs). Across exposure conditions, participants produced the same proportion of PDs regardless of whether they were speaking to Experimenter A (57.8%) or Experimenter B (55.0%; marginal main effect of Listener: $\chi^2(1) = 3.0849$, $p < .079$). Note that although this effect was marginally significant, the numerical pattern goes in the opposite direction as would demonstrate alignment – participants produced numerically more PDs to the DO-preferring experimenter than to the PD-preferring experimenter in the speaker-specific condition. Importantly, participants did not speak differently to the two experimenters as a function of exposure condition, producing the same

rate of PDs to their two partners regardless of whether the partners produced the identical syntactic distribution as each other or completely different distributions (no Syntax Exposure Condition × Listener interaction: $\chi^2(3) = 1.768$, $p < .622$).

The most important contrasts in this experiment were comparing PD production between the two single-structure exposure conditions (all-PD and all-DO) to test for global syntactic alignment, and between the two dual-structure exposure conditions (speaker-specific and mixed) to test for partner-specific syntactic alignment. Thus, planned sub-models were conducted to compare these conditions. In the single-structure comparison (all-PD vs. all-DO), participants produced significantly more PDs after being exposed to only PDs from both experimenters compared to only DOs from both experimenters (main effect of Syntax Exposure Condition: $\chi^2(1) = 5.721$, $p < .017$). This finding of global syntactic alignment shows that participants could modify their syntactic production within the constraints of this experimental procedure, and that hearing 24 sentences of a given syntactic structure in the present paradigm is enough to induce some form of syntactic alignment. Importantly, it shows that people both learn and align to syntactic statistical distributions aggregated across speakers, in a partner-independent manner. As expected, as both of a participant's partners produced the same syntactic distribution, participants produced the same syntactic distribution when speaking to each partner (no main effect of Listener: $\chi^2(1) = 1.825$, $p < .177$ or interaction between Syntax Exposure Condition and Listener: $\chi^2(1) = 0.903$, $p < .342$).

To test for the presence of partner-specific alignment, planned sub-models were conducted comparing the speaker-specific and mixed conditions. In both cases, participants heard a total of 24 PDs and 24 DOs during the exposure phase. If participants aligned in a partner-specific manner, participants in the speaker-specific condition should produce significantly more DOs to Experimenter A (who produced exclusively DOs) than to Experimenter B (who produced exclusively PDs), but participants in the mixed condition should produce the same rate of PDs to Experimenters A and B (who both produced both PDs and DOs). However, consistent with the prior experiments, there was no partner-specific alignment (no Syntax Exposure Condition × Listener interaction, $\chi^2(1) = 0.213$, $p < .644$). In the speaker-specific condition, participants produced 63.1% PDs to Experimenter A and 61.3% PDs to Experimenter B; in the mixed condition, participants produced 55.8% PDs to Experimenter A and 51.4% PDs to Experimenter B. There was no difference in overall PD production in the speaker-specific and mixed conditions (no main effect of Syntax Exposure Condition: $\chi^2(1) = 1.642$, $p < .200$), which is unsurprising given that participants received the same aggregate exposure to the two structures in the two conditions. Similarly, there was no main effect of Listener ($\chi^2(1) = 0.689$, $p < .407$).

Testing for an effect of partner-specific alignment at the finest level of granularity, in the speaker-specific condition alone, there was no difference in PD production to the two experimenters ($\chi^2(1) = 1.057$, $p < .304$). In fact, the slight numerical difference goes in the opposite direction as would indicate partner-specific syntactic alignment: Participants produced numerically fewer PDs to the PD-preferring listener than to the DO-preferring listener. As expected, given that syntactic exposure from the two experimenters was the same, there was no difference in PD production to the two experimenters in the mixed condition ($\chi^2(1) = 0.184$, $p < .668$).

Post-experiment fact test

When tested after the experiment, participants showed strong memory for which experimenter had said which fact. They correctly attributed the eight facts to their speakers significantly more often than chance (chance = 50%, mean = 97%, range: 75–100%; one-sample *t*(95) = 64.006, $p < .0001$).

See the Supplementary Materials for results from the other components of the Post-Experiment Questionnaire.

---

[2] For these pictures, many subjects produced a simple transitive sentence such as "A star is signing an autograph" or "The person is texting their best friend."
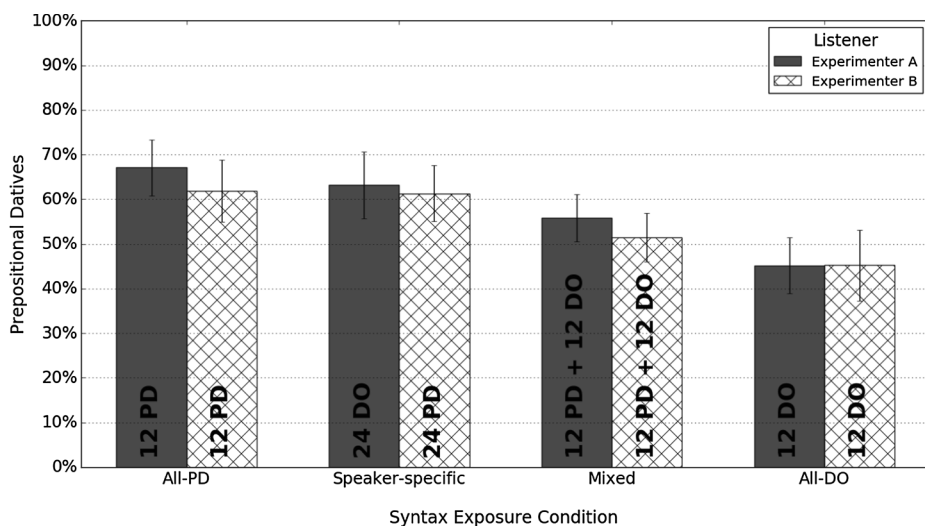
Fig. 5. Results from Experiment 4. Percentage of prepositional datives produced by participants, as a function of their listener's identity. The amount and type of dative exposure that a participant received from that experimenter is written on each bar. Error bars show standard error of the mean. PD = prepositional dative; DO = double object.

Discussion

The results from Experiment 4 replicated Experiments 1–3, and provided further evidence that speakers do not align their syntactic production to match their listener's demonstrated preferences, supporting an alignment mechanism which is sensitive to communicative utility. Participants were no more likely to produce PDs to a listener who *only* used PDs, compared to a listener who produced both structures equally often. In contrast, this experiment provided evidence for partner-independent alignment, as participants *did* align their syntactic production based on aggregated syntactic exposure, producing more PDs to both experimenters after being exposed to exclusively PDs, and more DOs to both experimenters after being exposed to exclusively DOs. This effect of global, partner-independent syntactic alignment demonstrates that hearing 24 sentences of a given structure is sufficient in the present paradigm to affect participants' syntactic learning and production, and that participants tracked the overall statistics of their syntactic exposure. Additionally, participants were demonstrably aware that they were interacting with two different experimenters who said different sentences, as they reliably attributed facts to the correct experimenter. Experiment 4 thus demonstrates that although speakers do not align their syntactic production in a partner-specific manner, they do learn the statistical distribution of the aggregated linguistic context and align partner-independently.

Syntax was chosen to investigate the mechanism driving alignment because in general, a shared conceptualization of syntax should not affect communicative ability. All of the employed structures are common in English and all interlocutors in the previous four experiments were highly fluent, native English speakers. As a result, it is a reasonable inference that both experimenters could understand both structures of an alternation, regardless of that speaker's distribution of produced sentences. But what about a situation in which partner-specific syntactic alignment could potentially improve communication? Experiment 5 tested whether speakers engage in partner-specific alignment when speaking with a non-native partner.

Experiment 5

The prior experiments demonstrated that alignment is not universally partner-specific; instead, speakers align partner-independently to aggregated linguistic statistics when doing so should not impair communicative utility. To further explore this hypothesis, Experiment 5 tested a linguistic context in which partner-specific syntactic alignment might actually help one's partner. Participants interacted with a non-native experimenter who might have been unfamiliar with the PD or

DO structure, more plausibly than the native experimenters.

There is prior evidence that native speakers modulate their production when speaking to non-native partners in order to reduce their listeners' processing difficulty. In syntax, they produce shorter and less syntactically and morphologically complex sentences, more sentences in canonical word order (e.g., SVO in English), include fewer clauses and adjectival, adverbial, and noun phrases per sentence, and reduce the number of complicated syntactic structures (Long, 1983). Native speakers also speak slower to non-native compared to native listeners (Scarborough, Brenier, Zhao, Hall-Lew, & Dmitrieva, 2007) and hyperarticulate vowels (Uther, Knoll, & Burnham, 2007). If the partner-specificity of alignment is modulated by communicative utility, then participants might syntactically align to the production statistics of a less proficient partner. If so, this would be the strongest evidence yet that speakers only align to a specific speaker's statistical distributions when doing so provides clear communicative utility, and align to the aggregated contextual statistical distribution otherwise. In contrast, if participants do not show partner-specific alignment to a non-native listener, perhaps lowered proficiency is not a strong enough cue for the communicative benefit of alignment, further supporting a model in which speakers align to just an aggregated statistical distribution across the current linguistic context, rather than individual distributions for each partner.

Method

Participants

Ninety-six students at the University of California, San Diego who did not take part in any of the previous experiments participated in this experiment for course credit. All reported being native, monolingual speakers of English.

Materials

The picture and sentence stimuli were the same as those used in Experiment 4. See Appendix A for a list of stimuli used in the experiment.

Procedure

The experimental design and procedure was identical to that of Experiment 4 except that one experimenter was a non-native English speaker with a heavy Mandarin accent. She began learning English at age 9 and did not live in an English-speaking country (USA) until age 18. The second experimenter was a native (unaccented) English speaker, as in all previous experiments. There were three people who acted as the native experimenter across participants, two female and

one male; different participants were tested by a different native experimenter (assigned non-systematically based on availability) paired with the same non-native experimenter. In the speaker-specific condition, syntactic preference of the native and non-native experimenter was counterbalanced across participants.

Results

Main experiment

Participants' descriptions were transcribed and coded offline (blind to syntax exposure condition and listener identity) as one of the two dative alternations (PD or DO), or neither. Descriptions that could not be categorized as one of the alternations, either because they were not a dative structure (28.6%) or were missing an argument (7.0%) were excluded. Seven pictures (out of 72) were excluded from the analyses because they had acceptable descriptions from fewer than 30% of participants.

The remaining sentence productions were submitted to a 4 (Syntax Exposure Condition: all-PD, speaker-specific, mixed, all-DO) × 2 (Listener Native-ness: native, non-native) GLMM. Participant means are reported below, and shown in Fig. 6.

Participants learned and aligned to the overall syntactic statistics (main effect of Syntax Exposure Condition: $\chi^2(3) = 16.941$, $p < .0007$). Participants produced the highest proportion of PDs after hearing only

PDs from both experimenters (all-PD exposure condition: 72.8% PDs), an intermediate amount after hearing both PDs and DOs (speaker-specific: 50.8% PDs; mixed: 54.1% PDs), and the lowest proportion of PDs after hearing only DOs from both experimenters (all-DO: 42.5% PDs). Across exposure conditions, participants produced the same proportion of PDs regardless of whether they were speaking to the native experimenter (55.6%) or the non-native experimenter (54.5%; no main effect of Listener Native-ness: $\chi^2(1) = 0.194$, $p < .660$), meaning that participants did not have an a priori expectation that the non-native experimenter would prefer one structure over the other. There was additionally no interaction ($\chi^2(3) = 0.705$, $p < .872$).

As in Experiment 4, planned sub-models were conducted to compare the single-structure conditions against each other (all-PD and all-DO) and the dual-structure conditions against each other (speaker-specific and mixed). Participants produced significantly more PDs after hearing only PDs from both experimenters (all-PD: 72.8% PDs), compared to hearing only DOs from both experimenters (all-DO: 42.5% PDs; $\chi^2(1) = 15.915$, $p < .0001$). Importantly, when restricted to the single-structure conditions, participants did not speak differently to the non-native experimenter when she produced the same distribution of syntactic structures as the native experimenter (no main effect of Listener Native-ness: $\chi^2(1) = 0.286$, $p < .593$ or interaction: $\chi^2(1) = 0.490$, $p < .484$). This finding of global syntactic alignment demonstrates that participants received sufficient exposure to modify their syntactic
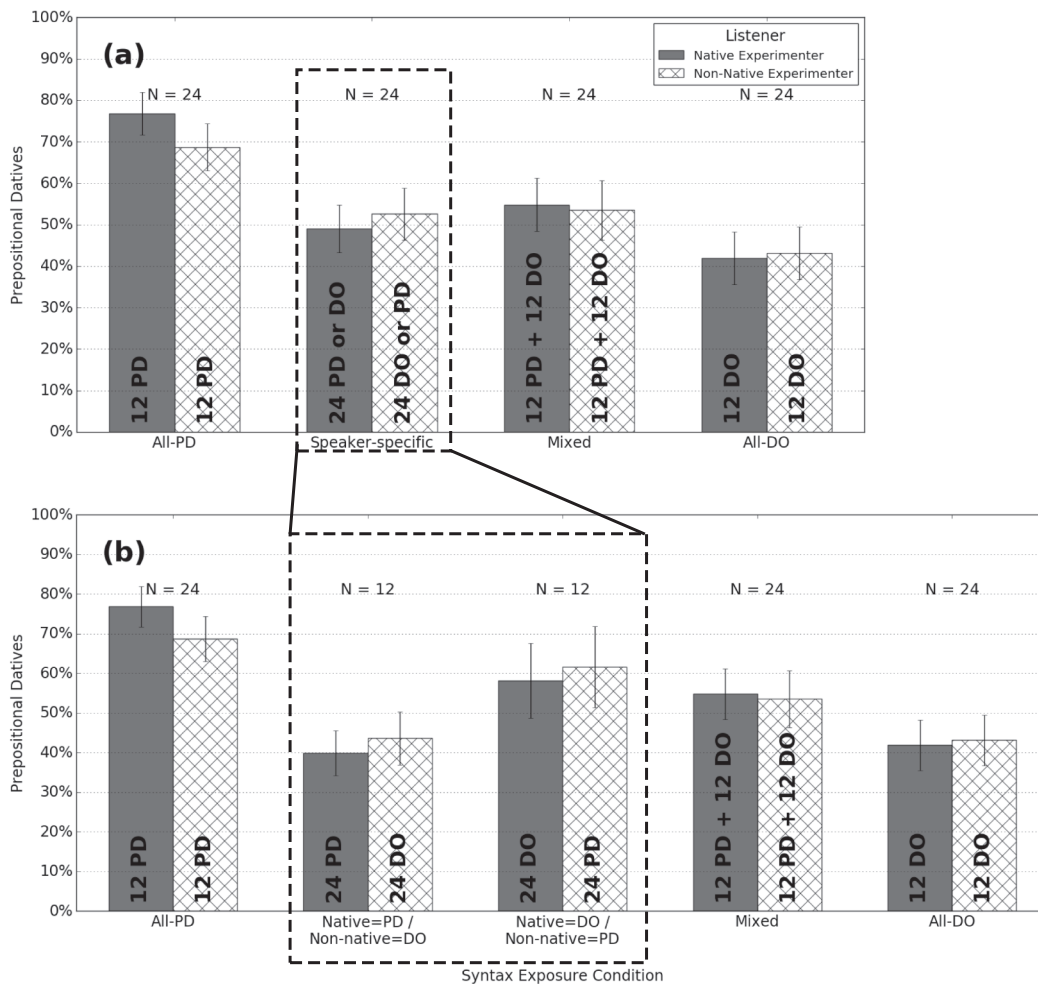

Fig. 6. (a) Results from Experiment 5. Percentage of prepositional datives produced by participants, as a function of whether they were speaking to the native or non-native experimenter. The amount and type of dative exposure that a participant received from that experimenter is written on each bar. (b) In the speaker-specific condition, half of the participants heard PDs from the native experimenter and DOs from the non-native experimenter; half heard the reverse. The all-PD, mixed, and all-DO conditions show the same data as in (a). Error bars show standard error of the mean.

production to both the native and non-native experimenter, and that participants learned (at least) the statistical distribution of their aggregated contextual input.

When restricted to the dual-structure conditions, there was no difference in overall PD production between the speaker-specific and mixed conditions (no main effect of Syntax Exposure Condition: $\chi^2(1) = 0.158$, $p < .691$). Critically, participants produced the same proportion of PDs to the two experimenters (no Listener Native-ness main effect or interaction; all $\chi^2(1) < 0.1$). Planned sub-models were conducted on PD production to the two experimenters within each exposure condition. Participants produced the same proportion of PDs to the native and non-native experimenter, both in the speaker-specific condition (to native: 49.0%; to non-native: 52.6%; $\chi^2(1) = 0.195$, $p < .659$) and in the mixed condition (to native: 54.8%; to non-native: 53.5%; $\chi^2(1) = 0.067$, $p < .795$).

However, unlike in Experiment 4, the two experimenters were not linguistically interchangeable. Thus syntactic exposure in the speaker-specific condition was counterbalanced across participants, such that half of the participants in this condition ($n = 12$) were exposed to exclusively PDs from the native experimenter and DOs from the non-native experimenter, and the other half of participants ($n = 12$) were exposed to the reverse mapping. (See Fig. 6 for the sub-groups' differing linguistic exposure.) Thus, if participants aligned partner-specifically, half should produce DOs and the other half produce PDs to the non-native experimenter, and combining the production data from these two sub-groups would average out the effects of partner-specific alignment. Therefore, to explore whether participants selectively aligned their production to the non-native experimenter when she produced a particular structure, a post hoc analysis was conducted on the speaker-specific condition using a 2 (Listener Native-ness: native, non-native) × 2 (Exposure Sub-group: native = PD/non-native = DO; native = DO/non-native = PD) GLMM to compare PD production in these two sub-conditions (see Fig. 6b).

Participants produced the same overall proportion of PDs when speaking to the native and the non-native experimenter (no main effect of Listener Native-ness: $\chi^2(1) = 0.080$, $p < .777$). There was a non-significant numerical trend of Exposure Sub-group ($\chi^2(1) = 2.548$, $p < .110$), as participants were numerically more likely to produce the structure that the non-native experimenter produced: The participants who heard DOs from the non-native experimenter produced fewer PDs to *both* experimenters (41.8% PDs); the participants who heard PDs from the non-native experimenter produced more PDs to *both* experimenters (59.9% PDs). However, even when interacting with a speaker who may have been less proficient, participants did not engage in partner-specific alignment, and did not speak differently to their two partners (no Listener Native-ness × Exposure Sub-group interaction: $\chi^2(1) = 0.183$, $p < .669$).

Post-experiment fact test

Participants were highly capable of ascribing specific statements to the correct experimenter. They correctly attributed an average of 95% of the facts to the experimenter who had mentioned them; this was significantly better than chance performance ($t(95) = 47.551$, $p < .0001$; chance = 50%; range: 50–100%).

See the Supplementary Materials for results from the other components of the Post-Experiment Questionnaire.

Discussion

Replicating the previous four experiments, participants did not align their syntactic production in a partner-specific manner, even when interacting with a potentially low-proficiency partner. However, they did align to the aggregated statistical input that they received. These results demonstrate that speakers are sensitive to variation in their overall syntactic input and can modify their production in response, when talking with both native and non-native partners.

This experiment provides the strongest evidence yet against alignment as a rote statistical process in which speakers track and align to all statistical regularities in the input. In the present experiment, not only did each partner produce a systematic and opposing syntactic statistical distribution, but one partner was a heavily-accented non-native speaker who might have provided a communicative utility cue to the participants. Nevertheless, even when interacting with two partners with differing levels of linguistic competence, participants still produced the same distribution of syntactic structures to each partner and aligned partner-independently.

General discussion

The current experiments investigated the statistical learning mechanism which underlies linguistic alignment, specifically by studying whether alignment remains partner-specific even in a linguistic situation where it does not add communicative utility. Across five experiments, alignment was not partner-specific: Participants did not modulate their own syntactic production to match their current listener's demonstrated syntactic preferences, and were just as likely to produce the structure that their current listener had *never* produced as to produce the structure that this listener had *always* produced. However, participants did align partner-independently to the aggregated statistics of the linguistic context, modulating their speech to the overall syntactic bias of their input.

The lack of partner-specific syntactic alignment occurred when participants were describing pictures for which they had already constructed an event representation (Experiments 1–3); pictures they had never seen or heard described (Experiments 2–3); and pictures for which they had contrastive evidence of their two partners' differing syntactic preferences and choice of one structure at the expense of another (Experiment 3). Participants also did not align their production to their partner's preference even when given substantially more exposure, enough to affect global syntactic production (Experiments 4–5). They did not even align uniquely to a partner for whom non-production of a particular structure could potentially be diagnostic of poor comprehension of that structure (Experiment 5). Together, these results demonstrate that partner-specific alignment is far from the default production strategy – that the language system is not just a rote statistical learning system which tracks, and aligns to, all aspects of contextual regularity that exist in the input. Rather, these results suggest that alignment is driven by a communicatively-sensitive mechanism which assesses the utility of aligning to different aspects of the linguistic context, and only aligning to those which provide a linguistic benefit.

However, the conclusion that speakers rely on partner-independent statistical distributions, rather than partner-specific ones, arises from null effects across the partner-specific conditions in the present experiments. The fact that the lack of partner-specificity was replicated across five experiments with different exposure conditions and participant samples strengthens the credibility of this conclusion. It is also important to note that these experiments are powerful enough to detect partner-specific syntactic alignment if it existed, as all experiments did show positive evidence for other types of syntactic alignment: event-specific alignment, when speakers' syntactic production was influenced by the structure used to describe a particular picture previously (Experiments 1–3) and partner-independent alignment, when speakers' syntactic production was influenced by the syntactic bias of the aggregate linguistic context (Experiments 4–5). This is a clear demonstration that participants represented the syntactic variability in the experiments, kept track of their syntactic input at an event-specific level as well as across the entire linguistic context, and modified their production in response to this input – they just did not modulate their syntactic production on a partner-by-partner basis.

In addition, to capitalize on the large number of participants that were run across these experiments, a cross-experiment analysis of the critical, speaker-specific alignment effect was run. This analysis included all trials from the above experiments in which participants were

exposed to two experimenters with contrasting and exclusive syntactic production, and in which participants described a picture that had not previously been described to them. (Thus it included the No Experimenter condition in Experiments 2 and 3 and Speaker-specific condition in Experiments 4 and 5. Note that data from Experiment 1 was not included in this analysis because all pictures were first described by one of the experimenters before the participant described them.) This analysis included 176 unique participants, 104 unique pictures, and 2412 trials across four experiments. However, there was still no support for partner-specific syntactic alignment, as, across experiments, participants were equally likely to produce a metric structure when speaking to the experimenter who always produced that structure (51.3% of trials) as when speaking to the experimenter who never produced that structure (50.9% of trials; no main effect of Listener: $\chi^2(1) = 0.056$, $p < .813$). Similarly, a Bayes factor was computed to quantify the strength of the evidence for the hypothesis that participants engaged in partner-specific syntactic alignment. The cross-experiment data as noted above produced a Bayes factor of 0.048 ($\pm 0.32\%$), which can be interpreted as "Strong evidence for H_0" (Jeffreys, 1961). This means the evidence prefers the model with no effect of speaker-specific alignment (only random effects) over the model that includes speaker-specific alignment (including the effect of Listener) by a factor of more than 20.

Another possible concern is that participants did not receive enough exposure to each speaker's syntactic preference to overwhelm their prior, partner-general distribution which includes knowledge of both structures. (In the extreme, it would be a poor conversational strategy to assume that a speaker is biased in favor of double object dative sentences if a listener only ever heard her produce one dative structure, which happened to be a DO.) However, Experiments 1–3 presented subjects with eight sentences of each event type from each experimenter, on par with the exposure in experiments that have shown partner-specific adaptation at other levels of language processing (e.g., Kraljic & Samuel, 2007; Metzing & Brennan, 2003; Trude & Brown-Schmidt, 2012). In Experiments 4–5, participants heard 24 sentences from each experimenter, which is considerably more than the amount of exposure in most other studies demonstrating some form of linguistic adaptation or alignment, including at the syntactic level (e.g., Fine et al., 2013; Kaschak, 2007; Kaschak et al., 2011; Kaschak et al., 2006; Kraljic, Brennan, & Samuel, 2008; Norris, McQueen, & Cutler, 2003). Additionally, the single-structure conditions in Experiments 4–5 demonstrated that hearing 24 sentences of a particular syntactic structure is indeed sufficient experience with the present stimuli and paradigm for participants to modify their syntactic productions and align (in a partner-independent way) to their input.

Additionally, participants were able to differentiate between their two partners and attribute specific information to the correct speaker. Participants displayed excellent memory for which experimenter had said which autobiographical fact in Experiments 4–5: Both experiments had an average accuracy rate well above 90%. Thus, participants were able to keep track of some partner-specific information within the experiments. Furthermore, the effect of event-specific alignment that was observed in Experiments 1–3 suggests that participants were able to remember (potentially implicitly) specific syntactic structures spoken by their interlocutors.

Models of syntactic priming and alignment

According to the Rational Speaker Model (Jaeger & Snider, 2013), partner-specific alignment should arise when the language processing system uses the partner-specific, as opposed to the partner-general, distribution to draw from when selecting which linguistic item will be produced. The most straightforward interpretation of the current results is that although speakers *learned* separate partner-specific distributions for each of their two partners (in addition to the aggregated, partner-general distribution of the overall linguistic environment), they did not reflect these

partner-specific distributions in their own production, likely because doing so does not add communicative utility. However, this presupposes that speakers tracked individual distributions for each of their partners, but without an explicit comprehension measure, it is possible that speakers did not. If so, this could be either because they *could* not track partner-specific statistics due to cognitive limitations, or, although partner-specific tracking is cognitively possible, they simply *did* not. The former possibility – that speakers are only able to track syntactic production aggregated across partners in their (recent) linguistic experience – is unlikely due to evidence from prior work in the adaptation literature. Although there is only sparse evidence for partner-specific learning in syntax in particular (Kamide, 2012; Squires, 2014), prior studies at other linguistic levels, including lexical entrainment (e.g., Brennan & Hanna, 2009; Brown-Schmidt, Yoon, & Ryskin, 2015; Metzing & Brennan, 2003; Yoon & Brown-Schmidt, 2014), phonemic perceptual learning (Kraljic & Samuel, 2007; Trude & Brown-Schmidt, 2012), and quantifier use (Yildirim, Degen, Tanenhaus, & Jaeger, 2016) demonstrate that people can easily detect and track separate linguistic distributions tied to specific partners, and, as noted in the prior section, participants in the present experiments heard equivalent or substantially more instances from each partner than seems to be necessary to do so. The latter possibility – that speakers could have but did not track partner-specific distributions – has similar implications for the mechanism underlying alignment as if speakers did learn partner-specifically, but nevertheless drew from the partner-general distribution for their own production. As speakers are demonstrably able to learn partner-specific distributions at other linguistic levels (e.g., lexical, phonemic), there must be a reason why this would not happen for syntax. Communicative utility as a decision mechanism still applies – whether it influences whether speakers even *learn* partner-specific distributions, or do learn them but do not draw upon them for production. Further research should explicitly investigate partner-specific learning in comprehension in the case of communicative non-necessity, to determine whether such distributions are learned but not used, or are not even learned in the first place.

It is of course also possible that the framework suggested by the Rational Speaker Model is not correct. Other models of syntactic priming generally cannot make explicit predictions of either partner-specificity or partner-independence of alignment because, as far as we know, none include parameters about priming as modulated by the identity of the listener or communicative utility. However, different models propose different mechanisms to underlie syntactic priming, which are relatively more or less practical to drive alignment, and thus could predict different patterns of data. Pickering and Branigan (1998) propose that syntactic priming comes about as the result of transient activation at the level of the individual sentence, in which exposure to a particular structure boosts the activation level of that structure, making it more accessible and thus more likely to be produced in upcoming discourse. This mechanism is unlikely to be compatible with partner-specific alignment, as it is hard to see how structure-specific transient activation could be tied with long-term learning to a particular speaker. Thus this model should predict a strong recency effect, in which more recently-heard structures are more likely to be produced globally, but irrespective of the listener's identity, and always partner-independent alignment, independent of communicative necessity. This model is compatible with the current results: Participants who received structure-biased input (all-PD and all-DO in Experiments 4 and 5) produced more of that structure, which could have been due to accumulating transient activation boosts over the exposure phases. In contrast, participants who received structure-balanced input (Experiments 1–3 and speaker-specific and mixed conditions in Experiments 4 and 5) should have the transient activation for each structure cancelled out by the opposing alternation shortly thereafter, and thus show no overall bias in favor of their current partner's preference. However, it is hard to square this model's predictions with prior demonstrations of partner-specific adaptation and alignment, though those have largely been confined to other linguistic levels, rather than syntax.

Several other models argue (e.g., Bock & Griffin, 2000; Chang, Dell, & Bock, 2006; Reitter, Keller, & Moore, 2011) that syntactic priming is

caused by longer-term, implicit learning, in which linguistic statistics are learned over time and aggregated experience. Reiter and colleagues' (2011) model uses associative learning of co-occurrences between a syntactic structure and lexical-semantic material such that future use of that semantic material becomes a retrieval cue for the syntactic structure (the lexical boost effect). A similar mechanism could learn the co-occurrence between a syntactic structure and the current speaker to produce partner-specific alignment, with the speaker serving as a retrieval cue for their preferred structure. However, perhaps the speaker's identity is not as strong of a retrieval cue as a lexical item, explaining why syntactic production shows the lexical boost effect but not partner-specific alignment.

Production diagnosticity as a mechanism for partner-specific alignment

The question of what statistics people align to from their input, as measured by what they reflect in their own production – statistical distributions tied to individual speakers, or an overall distribution across the context – seems to clearly fall on the side of aligning to the overall linguistic context. This in turn supports a mechanistic account of linguistic alignment which takes communicative utility into account, in order to explain both partner-specific effects in certain circumstances (as with lexical entrainment) as well as partner-independent effects in others (as in the present syntactic experiments). Why should the specificity of alignment depend on its perceived communicative utility? One possibility is in regards to a speaker's perception of his partner's comprehension ability. Each speaker produces a variable set of linguistic features which differs, to some extent, from those produced by other speakers. A person's linguistic system needs to determine which of these idiosyncratic features signal something meaningful about his interlocutor's language processing ability and which are a byproduct of unimportant variability. Partner-specific alignment may occur only in the former case, to the degree that an interlocutor's production signals their level of comprehension.

There is some suggestion from prior work that speakers employ this production diagnosticity. Branigan, Pickering, Pearson, McLean, and Brown (2011) found that people showed more lexical alignment to object labels which were allegedly produced by a computer as opposed to another human, an effect which was heightened when interacting with a more "basic" as compared to "advanced" computer. Although these results do not speak to the ability to learn a partner's overall linguistic distribution and align in a partner-specific way, they do suggest that people may be sensitive to their partners' varying levels of linguistic competency and can modulate their own production to take that into account. Specifically, these results suggest that speakers are more likely to mimic their partner's linguistic choices when their partner might have low linguistic competency.

As discussed in the introduction, an important feature of many experiments that demonstrate lexical entrainment is that the objects being described are somehow unusual and require either inventing a new referring word or phrase, or using a subordinate-level category word in violation of the basic-level preference (Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986; Rosch, 1978), and the tacit lexical agreement between interlocutors is based on a shared conceptualization of the object. One way for a speaker to know that his partner has agreed with his proposed conceptual pact is if the partner produces the same label for a particular object. Thus, partner-specific alignment (in this example, lexical entrainment) occurs because the partner's production (the agreed-upon label) is diagnostic of her own comprehension expectation of how this object will be referenced. More generally, when a speaker expects that a partner more easily understands what the partner herself produced, the speaker should be biased to produce language with the same features as his partner in order to facilitate communication. This is manifest as partner-specific alignment.

In contrast, the linguistic materials used in the present experiments were all very common syntactic structures. As noted above, the syntactic alternations used in the present experiments are learned by native

speakers at a very young age (Shimpi et al., 2007; Snyder & Stromswold, 1997) and thus it is a reasonable inference that a fluent, adult conversational partner should be familiar with all of the syntactic alternations used here. Thus participants could have had a strong prior expectation that both of their partners, even if they each only produced one of the two structures in an alternation, could nevertheless successfully comprehend both and thus partner-specific alignment should not have a big impact on comprehensibility. This allowed us to set up a linguistic environment to test whether indeed speakers modulate the linguistic statistics they align to as a function of communicative utility, or align partner-specifically by default.

This explanation has some support from a previous study on listener-directed syntactic adaptation (Valian & Wales, 1976). When reading aloud, speakers reformulated sentences to have clearer and more explicit syntactic relations after a listener expressed confusion. However, this happened almost exclusively when the original sentence was a (somewhat) complex structure and there was a more explicit one available – for example, revising a subject or object relative to include the complementizer "that." Speakers rarely revised their original structure when it was already the more explicit version of an alternation, even if the listener asked for clarification. This result suggests that speakers may be willing to modify their syntactic production based on their listener, but only if such adaptation is warranted by potential listener confusion.

Conclusion

The experiments presented here show that linguistic alignment, at least within the domain of syntax, is a partner-independent process and not a partner-specific one. This suggests that alignment is driven not by a rote statistical mechanism which aligns to all aspects of contextual variability, but rather a utility-sensitive mechanism, which aligns to partner-specific statistical distributions to the extent that they improve communication. In the case of lexical entrainment or phonetic adaption, where individual production differences affect how words or sounds are comprehended, learning and aligning to individual statistical distributions has utility and thus the system does so. In contrast, in the case of syntactic alignment, the difference between partners' productions is not an important source of variability because it likely does not provide diagnostic information about comprehension ability, and thus partner-specific alignment would not enhance communicative utility. As a result, speakers align to the partner-general syntactic distribution, which models the statistics of the overall context, and do not align to the individual syntactic distributions of each partner. Thus, alignment is driven by a communicatively-informed, and not just statistically-determined, mechanism.

Declaration of Competing Interest

The authors declare they have no competing interests.

Appendix A. Stimuli used in Experiments 1–5

Sentence (Active, On, or PD structure)	Exp 1	Exp 2	Exp 3	Exps 4 & 5
Transitive Events				
A beachball is hitting a car	x	x	x	
A bee is stinging a man	x	x	x	
A brown horse is chasing a white horse	x	x	x	
A car is popping a balloon	x	x		
A clown is pulling a dog in a wagon	x	x	x	
A flyswatter is swatting a fly	x	x	x	
A girl is pulling a horse	x	x	x	
A man is kidnapping a woman	x	x	x	
A tornado is destroying a barn	x	x	x	
A whale is swallowing a man	x	x	x	
A wine bottle is cutting a girl	x	x		
A wrecking ball is destroying a building	x	x	x	
Lightning is striking a church	x	x	x	
Lightning is striking a golfer	x	x	x	
The cowboy is branding the horse		x		
The devil is poking an angel	x	x	x	
The doctor and patient are studying the x-ray		x		
The elephant is swinging the branch		x		
The girl is hanging the letters		x		
The jack in the box is scaring a little girl	x	x	x	
The missile is hitting the plane		x	x	
The mother is holding the baby		x		
The police woman is issuing a ticket		x		
The tree is smashing the roof[a]		x	x	
Locative Events				
A chef is brushing butter on a turkey	x	x	x	
A creepy guy is sticking needles in a voodoo doll	x	x	x	
A girl is smearing paint on her face	x	x	x	
A girl is spreading jam on her toast	x	x		
A man is cramming clothes into his drawers	x	x	x	
A man is packing beers in a cooler	x	x	x	
A man is smearing shaving cream on his face	x	x		
A man is splattering paint on a canvas	x	x	x	
A man is spraying insecticide on his plants	x	x	x	
A man is stuffing marshmallows in his mouth	x	x	x	
A woman is rubbing lotion on a baby	x	x		
A woman is spritzing water on her plant	x	x	x	
An elephant is spraying water on a clown	x	x	x	
Men are loading toxic material on a truck	x	x	x	
Santa Claus is stuffing fruit into the stockings	x	x	x	
Soldiers are draping a flag over a coffin	x	x	x	
The clown is packing shoes in his suitcase		x		
The cowboy is branding a heart on a horse			x	
The farmer is spreading seeds on the field		x		
The girl is sprinkling salt into the soup		x		
The man is drizzling nacho cheese on the nachos		x		
The man is stacking cups on the tray		x	x	
The man is stocking canned food in the bomb shelter		x	x	
The mother is dabbing alcohol on the little girl's cut		x		
The nurse is wrapping a bandage on the patient's arm		x		
Dative Events				
A baker is showing a wedding cake to the bride				x
A ballerina showing her tiara to a girl				x
A bank teller is handing money to a robber				x
A boy is handing flowers to a girl				x
A boy is throwing a backpack to another boy				x
A boy is tossing the ball to his dad	x	x		x
A car salesman is selling a car to a happy man				x
A chef is tossing a tomato to his assistant				x
A climber is tossing the rope to his friend				x
A coach is showing the gameplan to the players				x
A doctor is handing pills to a patient[b]	x	x	x	x
A girl is handing a paintbrush to a man	x	x	x	x
A girl is mailing a letter to Santa				x
A girl is texting a picture to her friend				x
A girl is throwing a bone to a dog				x
A girl is tossing a leash to the dog	x	x		x
A grandma is offering cookies to her grandchildren				x
A guy is showing his ID to the bouncer[c]	x	x	x	x

Event				
A lawyer is showing evidence to the judge	x	x	x	x
A man is giving a gift to a little girl	x	x	x	x
A man is giving a hat to a clown	x	x	x	x
A man is showing his sports car to another man				x
A nurse is giving a lollipop to a boy				x
A nurse is giving crutches to a girl	x	x	x	x
A policeman is issuing a ticket to a driver				x
A rockstar is giving his autograph to a fan				x
A runner is passing the baton to his teammate				x
A sailor is pushing a crate to another sailor				x
A stewardess is serving coffee to a man	x	x	x	x
A student is passing a note to her friend				x
A teacher is giving the exam to a student				x
A teenaged girl is showing her tattoo to a motorcyclist				x
A UPS driver is bringing a package to a man				x
A waiter is serving steaks to the men				x
A waitress is offering drinks to the people	x	x		x
A woman is handing a dollar bill to a man				x
A woman is offering a ball of yarn to a cat				x
A woman is showing a dress to a man	x	x	x	x
A woman is showing the blueprint to the construction worker	x	x	x	x
A woman is throwing a beachball to a patient	x	x		x
An artist is handing her artwork to a woman				x
An old woman is feeding breadcrumbs to some pigeons				x
An old woman is giving a birthday cake to an old man	x	x	x	x
An Olympian is showing his medal to his family				x
One basketball player is passing the ball to another player				x
One guy is throwing the frisbee to the other guy	x	x		x
One juggler is throwing a club to another juggler				x
One monkey is throwing a banana to another monkey				x
The actress is reading her lines to the director				x
The baby sitter is bringing the crying baby to the mother				x
The ball boy is rolling a tennis ball to the tennis player				x
The boy is feeding a cracker to his parrot				x
The boy is handing a letter to the girl			x	x
The cafeteria worker is giving a tray to the woman		x		x
The jockey is feeding a carrot to his horse				x
The librarian is lending a book to the boy		x		x
The lifeguard is tossing a rescue buoy to the swimmer				x
The mama bird is feeding a worm to the baby bird				x
The man is giving a bowl of milk to his cat				x
The man is handing the plate of dog food to his wife[d]		x	x	x
The mother is handing a cup of water to her daughter		x	x	x
The nun is offering a candy to the boy		x		
The oil tycoon is paying a bribe to the senator				x
The patient is giving her prescription to the pharmacist		x	x	x
The professor is giving a diploma to a student				x
The quarterback is throwing the football to the receiver				x
The realtor is showing a house to an elderly couple				x
The slave is offering grapes to the pharaoh				x
The soccer player is passing the ball to her teammate				x
The teacher is reading a story to the class		x		x
The tour-guide is showing the statue to the tourists				x
The waitress is handing a menu to the man		x	x	x
The woman is offering her jacket to a homeless man				x

Intransitive Events

Event				
A boy is sliding down the slide				x
A girl in a raincoat is looking out the window				x
A girl is peering into the cookie jar				x
A homeless man is begging for money				x
A man in a winter hat is walking through the woods				x
A tree is falling on the house				x
The baby is crawling				x
The boy is roller skating				x
The boy is somersaulting				x
The boy is yawning				x
The faucet is running				x
The house is on fire				x
The little girl is crying really hard				x
The man is diving into a swimming pool				x
The man is sitting in a chair				x
The man is sweating from the heat of the sun				x
The man is swimming in the ocean				x
The owl is flying				x
The pot is boiling				x
The volcano is erupting				x

The woman is sleeping	x
The woman is sneezing	x
The woman is sunbathing on the beach	x
Two horses are running	x

[a] Experiment 3 presented the sentence: "The tree is destroying the roof."
[b] Experiments 3, 4, and 5 presented the sentence "A nurse is handing pills to a patient."
[c] Experiments 1, 2, and 3 presented the sentence "Some guy is showing his ID to the bouncer."
[d] Experiment 2 presented the sentence "The mustachioed man is handing the plate of dog food to his wife."

Appendix B. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jml.2019.104037.

References

Atkinson, J. E. (1976). Inter-and intraspeaker variability in fundamental voice frequency. *The Journal of the Acoustical Society of America, 60*, 440–445.

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics, 40*, 177–189.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48.

Bergmann, K., & Kopp, S. (2012). *Gestural Alignment in Natural Dialogue* (pp. 1326–1331). Austin, TX: Cognitive Science Society.

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

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75*, B13–B25.

Branigan, H. P., Pickering, M. J., McLean, J. F., & Cleland, A. A. (2007). Syntactic alignment and participant role in dialogue. *Cognition, 104*, 163–197.

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*, 41–57.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1482–1493.

Brennan, S. E., & Hanna, J. E. (2009). Partner-specific adaptation in dialog. *Topics in Cognitive Science, 1*, 274–291.

Brown-Schmidt, S., Yoon, S. O., & Ryskin, R. A. (2015). People as contexts in conversation. *Psychology of Learning and Motivation, 62*, 59–99.

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

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1–39.

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.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria <https://www.R-project.org/>.

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE, 8*(10), 1–18.

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

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*, 251–279.

Gregory, S. W., Dagan, K., & Webster, S. (1997). Evaluating the relation of vocal accommodation in conversation partners' fundamental frequencies to perceptions of communication quality. *Journal of Nonverbal Behavior, 21*, 23–43.

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

Gruberg, N., Ostrand, R., Momma, S., & Ferreira, V. S. (2019). Syntactic entrainment: The repetition of syntactic structures in event descriptions. *Journal of Memory and Language, 107*, 216–232.

Hanson, H. M., & Chuang, E. S. (1999). Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *The Journal of the Acoustical Society of America, 106*, 1064–1077.

Haywood, S. L., Pickering, M. J., & Branigan, H. P. (2005). Do speakers avoid ambiguities during dialogue? *Psychological Science, 16*, 362–366.

Holler, J., & Wilkin, K. (2011). Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of Nonverbal Behavior, 35*, 133–153.

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition, 127*, 57–83.

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). New York: Oxford University Press.

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics, 34*, 485–499.

Kamide, Y. (2012). Learning individual talkers' structural preferences. *Cognition, 124*,

Kaschak, M. (2007). Long-term structural priming affects subsequent patterns of language production. *Memory and Cognition, 35*, 925–937.

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*, 1133–1139.

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

Kimbara, I. (2008). Gesture form convergence in joint description. *Journal of Nonverbal Behavior, 32*, 123–131.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review, 122*, 148–203.

Kraljic, T., Brennan, S. E., & Samuel, A. G. (2008). Accommodating variation: Dialects, idiolects, and speech processing. *Cognition, 107*, 54–81.

Kraljic, T., & Samuel, A. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language, 56*, 1–15.

Lawless, H. T. (1984). Flavor description of white wine by "Expert" and nonexpert wine consumers. *Journal of Food Science, 49*, 120–123.

Levelt, W. J., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology, 14*, 78–106.

Liu, L., Burchill, Z., Tanenhaus, M. K., & Jaeger, T. F. (2017). *Failure to replicate talker-specific syntactic adaptation* (pp. 2616–2621). Austin, TX: Cognitive Science Society.

Long, M. H. (1983). Linguistic and conversational adjustments to non-native speakers. *Studies in Second Language Acquisition, 5*, 177–193.

Matarazzo, J. D., Weitman, M., Saslow, G., & Wiens, A. N. (1963). Interviewer influence on durations of interviewee speech. *Journal of Verbal Learning and Verbal Behavior, 1*, 451–458.

Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language, 49*, 201–213.

Miller, J. L., Grosjean, F., & Lomanto, C. (1984). Articulation rate and its variability in spontaneous speech: A reanalysis and some implications. *Phonetica, 41*, 215–225.

Mol, L., Bogers, R., & Bouwens, T. (2012). *Automated and Partner-Specific Factors Influencing Lexical Entrainment* (pp. 755–760). Austin, TX: Cognitive Science Society.

Namy, L. L., Nygaard, L. C., & Sauerteig, D. (2002). Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology, 21*, 422–432.

Natale, M. (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology, 32*, 790–804.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*, 204–238.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America, 119*, 2382–2393.

Pardo, J. S. (2010). Expressing Oneself in Conversational Interaction. In E. Morsella (Ed.). *Expressing Oneself/Expressing One's Self: Communication, Cognition, Language, and Identity* (pp. 183–196). New York: Psychology Press/Taylor & Francis.

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*, 427–459.

Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science, 35*, 587–637.

Reitter, D., & Moore, J. D. (2014). Alignment and task success in spoken dialogue. *Journal of Memory and Language, 76*, 29–46.

Rosch, E. (1978). Principles of categorization. In E. Rosch, & B. B. Lloyd (Eds.). *Cognition and categorization*. Hillsdale, NJ: Erlbaum.

Scarborough, R., Brenier, J., Zhao, Y., Hall-Lew, L., & Dmitrieva, O. (2007). An acoustic study of real and imagined foreigner-directed speech. In J. Trouvain & W. J. Barry (Eds.), Proceedings of The International Congress of Phonetic Sciences (pp. 2165–2168). Saarbrücken, Germany.

Shimpi, P. M., Gámez, P. B., Huttenlocher, J., & Vasilyeva, M. (2007). Syntactic priming in 3-and 4-year-old children: Evidence for abstract representations of transitive and dative forms. *Developmental Psychology, 43*, 1334–1346.

Shockley, K., Sabadini, L., & Fowler, C. A. (2004). Imitation in shadowing words. *Perception & Psychophysics, 66*, 422–429.

Snyder, W., & Stromswold, K. (1997). The structure and acquisition of English dative constructions. *Linguistic Inquiry, 28*, 281–317.

Squires, L. (2014). Talker specificity and the perception of grammatical variation. *Language, Cognition and Neuroscience, 29*, 856–876.

Street, R. L. (1984). Speech convergence and speech evaluation in fact-finding interviews. *Human Communication Research, 11*, 139–169.

ten Bosch, L., Oostdijk, N., & de Ruiter, J. P. (2004). Durational aspects of turn-taking in spontaneous face-to-face and telephone dialogues. In P. Sojka, I. Kopecek, & K. Pala (Eds.). *Text, Speech and Dialogue: Proceedings of the 7th International Conference TSD 2004* (pp. 563–570). Heidelberg: Springer.

Terrell, T. D. (1976). Some theoretical considerations on the merger of the low vowel phonemes in American English. Proceedings of the Second Annual Meeting of the Berkeley Linguistics Society, Department of Linguistics (pp. 350–359). Berkeley, California: University of California at Berkeley.

Trude, A. M., & Brown-Schmidt, S. (2012). Talker-specific perceptual adaptation during online speech perception. *Language and Cognitive Processes, 27*, 979–1001.

Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication, 49*, 2–7.

Valian, V., & Wales, R. (1976). What's what: Talkers help listeners hear and understand by clarifying sentential relations. *Cognition, 4*, 155–176.

Webb, J. T. (1969). Subject speech rates as a function of interviewer behaviour. *Language and Speech, 12*, 54–67.

Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language, 31*, 183–194.

Yildirim, I., Degen, J., Tanenhaus, M. K., & Jaeger, T. F. (2016). Talker-specificity and adaptation in quantifier interpretation. *Journal of Memory and Language, 87*, 128–143.

Yoon, S. O., & Brown-Schmidt, S. (2014). Adjusting conceptual pacts in three-party conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 919–937.