

2013

Attribution of Mutual Understanding

Carl Vogel, Ph.D.

Follow this and additional works at: <http://brooklynworks.brooklaw.edu/jlp>

Recommended Citation

Carl Vogel, Ph.D., *Attribution of Mutual Understanding*, 21 J. L. & Pol'y ().

Available at: <http://brooklynworks.brooklaw.edu/jlp/vol21/iss2/6>

This Article is brought to you for free and open access by BrooklynWorks. It has been accepted for inclusion in Journal of Law and Policy by an authorized administrator of BrooklynWorks. For more information, please contact matilda.garrido@brooklaw.edu.

ATTRIBUTION OF MUTUAL UNDERSTANDING

*Carl Vogel**

INTRODUCTION

This article explores advances in a method of analysis of conversational interaction, as recorded through text transcripts, for evidence of grounding in order to quantify certainty of mutual understanding. It is necessary to take into account aspects of communication in which certainty of having arrived at a common understanding of dialogue content must be pessimistically assessed. It may be that in many or even most contexts, the urgency of the linguistic elements of communication is negligible. It is a relatively rare event for linguistic acts, independently of other forms of communication, to have a distinctive, measurable, impact on human survival, and therefore, where such events exist, they tend to be spectacular.¹ That a failure to achieve mutual understanding does not typically result in catastrophic events does not constitute

* Computational Linguistics Group, O'Reilly Institute, Trinity College Dublin, Dublin 2, Ireland (vogel@tcd.ie). Thanks are due to Lawrence Solan for organizing the Brooklyn Law School Authorship Attribution Workshop and his acknowledgement of the breadth of problems that fall into the category of linguistic attribution in legal contexts. Additionally, I am very grateful to Carole Chaski for discussions of this material and for making me aware of the work of Linda Smith.

¹ See, for example, the discussion of Flying Tiger Flight 66, on February 19, 1989, in which the air traffic control directive, “descend two four zero zero” was evidently understood as having a preposition, “to” rather than a numeral “two,” with the consequence that the aircraft flying at an altitude of 400 feet crashed into a hill 437 feet above sea level, killing all four people on the aircraft. STEPHEN CUSHING, FATAL WORDS: COMMUNICATIONS CLASHES AND AIRCRAFT CRASHES 14 (1994).

sufficient evidence to suggest that conversation is generally successful. It appears to be largely a matter of doctrine whether communicators are assumed to generally understand each other and signify the same ideas with the same language—this is known as the doctrine of intersubjective conformity.² The skeptical position that neither interlocutors (the participants in dialogues) nor external observers can ever completely verify whether the interlocutors have really understood each other is so robust that this skeptical position cannot be refuted. However, a weaker notion of intersubjective conformity is available. The weaker notion is that dialogue participants or observers may pragmatically behave as if there has been mutual understanding unless contrary evidence emerges. Nonetheless, in some contexts, forensic ones in particular, it seems safer to adopt the null hypothesis about communicative success which holds that communication attempts have been unsuccessful unless positive evidence exists that mutual understanding has emerged sufficiently to make one reject the null hypothesis and accept the alternative hypothesis that communication has been successful, rather than to assume success by default.

This article focuses on the development and testing of objective measures for assessing the likelihood of shared understanding of linguistic communication in contexts where shared understanding has a critical role, such as in forensic interrogations or other courtroom interactions. After motivating a principle of skepticism in assessing the likelihood of mutual understanding emerging for all participants in any given dialogue, a method is described which is deployed here to quantify levels of engaged interaction as a proxy measure for mutual understanding. The basic idea is that where positive evidence is needed to assert that dialogue participants have understood each other, levels of interaction that are statistically significant in divergence from random interactions provide a tangible basis for asserting that corresponding levels of mutual understanding may have been achieved. The method is illustrated with respect to an excerpt of a dialogue transcript

² TALBOT J. TAYLOR, MUTUAL MISUNDERSTANDING: SCEPTICISM AND THE THEORIZING OF LANGUAGE AND INTERPRETATION 29 (1992).

about which it is relatively easy for third party observers to form opinions about the communicative success achieved, then the method is applied to two transcripts available in public records of cases in which levels of mutual understanding has been contested. It is argued that the method of analysis is able to contribute useful facts to debates about the level of mutual understanding achieved in dialogues in which that form of communicative success matters.

Many contexts of legal interpretation are primed by principles associated with criminal trials, rigorously applied (i.e., “presumed innocent, unless proven guilty”). Presently, it is argued that the null hypothesis regarding the success of linguistic communication is that language use is ineffective unless proven effective.³ This statement is jarring on first encounter because language use in communication is largely taken for granted as being as effective as the use of language in thought. However, one need only reflect on the many sorts of ambiguity that exist in language (i.e., sonic, syntactic, semantic) as well as their potential for combinatoric increase in the number of potential meanings to realize how great the chances are for miscommunication to arise through linguistic channels.⁴ Indeed, much literature about theories, models, and simulations

³ See *id.* at 30. Taylor presents such arguments at a meta level, in relation to possible rebuttals and resolutions; here, the proposal is to make do with this skeptical position rather than to argue against it. See generally *id.*

⁴ Consider an example:

Suppose a sentence has three ambiguous lexical items and two (disjoint) places with attachment ambiguities; even if each ambiguity allows only two possibilities, the sentence will have, in principle, $2^5 = 32$ interpretations. A simple example satisfying this description is given in (7); others would be easy to construct.

7. Old friends and acquaintances remembered Pat’s last time in California.

Here old can mean aged or long-term (or former) and can modify either friends and acquaintances or just friends; last can mean final or previous; time can mean occurrence or duration (e.g. if Pat was a racer), and in California can modify remembered or time.

Thomas Wasow et al., *The Puzzle of Ambiguity*, in MORPHOLOGY AND THE WEB OF GRAMMAR: ESSAYS IN MEMORY OF STEVEN G. LAPOINTE (C. Orhan Orgun & Peter Sells eds., 1998).

of language evolution explores the emergence of linguistic systems from the assumption that the first communicators began with shared thoughts and built language on prior, shared, thought.⁵ When the context of communication is legal, given its conventional location of burden of proof, it can be more important to attempt to quantify the level of understanding that could have been achieved by linguistic means, for example, during testimony, particularly when nonnative speakers of a language are involved.⁶

Consider the following transcript of a courtroom dialogue⁷:

⁵ Of course, this bootstrapped semantic infallibility, or “telepathy” approach is not universal in the language evolution literature. *See, e.g.*, Andrew D.M. Smith, *The Inferential Transmission of Language*, 13 ADAPTIVE BEHAV. 311, 311–23 (2005); Martin Bachwerk & Carl Vogel, *Establishing Linguistic Conventions in Task-Oriented Primeval Dialogue*, PROC. COST 2102 INT’L CONF., 2010, at 48–55; Martin Backwerk & Carl Vogel, *Language and Friendships: A Co-Evolution Model of Social and Linguistic Conventions*, 9TH INT’L CONF. ON EVOLUTION LANGUAGE, 2012, at 34–41; James R. Hurford, *Biological Evolution of the Saussurean Sign as a Component of the Language Acquisition Device*, 77 LINGUA 187, 187–222 (1989) (analyzing the biological evolution of the Saussurean sign as a component of the language acquisition device); James R. Hurford & Simon Kirby, *Co-Evolution of Language-Size and the Critical Period*, in SECOND LANGUAGE ACQUISITION AND THE CRITICAL PERIOD HYPOTHESIS 39, 39–63 (David Birdsong ed., 1999); James Hurford, *The Evolution of the Critical Period for Language Acquisition*, 40 COGNITION 159, 159–201 (1991); Simon Kirby, *Syntax Without Natural Selection: How Compositionality Emerges from Vocabulary in a Population of Learners*, in THE EVOLUTIONARY EMERGENCE OF LANGUAGE: SOCIAL FUNCTIONS AND THE ORIGINS OF LINGUISTIC FORM 303, 303–23 (James Hurford et al. eds., 2000).

⁶ The burden of proof is not identically located in all legal contexts. In contract law, it is typical to presume that an individual understands the contractual issues, including the fine print; for example in relation to arguing the invalidity of a contract on the basis that it presupposes mistaken assumptions, “[t]he basic rule is that a mistaken assumption must be shared and be of fundamental importance if it is to be the basis for setting aside an otherwise valid contract. This is rarely satisfied.” STEPHEN A. SMITH, CONTRACT THEORY 283 (1993). Extra measures are needed to establish the invalidity of a contract due to deliberate drafting obfuscation, but this can be achieved. Melvin A. Eisenberg, *The Limits of Cognition and the Limits of Contract*, 47 STAN. L. REV. 211, 240–49 (1995).

⁷ *People v. Herrero*, 756 N.E.2d 234, 242 (Ill. App. Ct. 2001).

- BASTOUNES Mr. Herrero is a Spanish speaking individual. We ordered an interpreter and he isn't here. He understands well enough that if you want you can admonish him on the record. He understands well enough what is going on in terms of picking this jury. He doesn't have a problem and wants to proceed this way. Perhaps we should put that on the record and tomorrow morning when we do opening statements and evidence
- COURT Mr. Herrero would you step up here, please.
- BASTOUNES I'm for sure that we would have an interpreter.
- BRODE We would like to put on the record that neither one of these individuals ever needed an interpreter.
- BASTOUNES I think the first time I was here with Mr. Herrero at the bond hearing didn't we? I just wanted to be sure that the record is clear and that there is no alleged error later on we will want an interpreter for the trial and it should be no problem getting one tomorrow.
- BASTOUNES For the record I have discussed with my client Mr. Herrero his desire to proceed this afternoon with picking the jury and he has indicated to me that he understands and wish well enough for that portion of the trial and wants to proceed.
- BASTOUNES Mr. Herrero, is it your desire now to proceed with picking the jury?
- HERRERO Yes.
- BASTOUNES Without an interpreter?
- HERRERO Yes.

- BASTOUNES Do you understand what I'm saying to you now, is that correct?
- HERRERO Yes.
- BASTOUNES Okay. Judge, if you want to inquire further.
* * *
- COURT Mr. Herrero, have you understood the conversation that has taken place in the last ten minutes or so?
- HERRERO I understand a little bit.
- COURT Mr. Herrero, do you have any objection to picking the jury now without the interpreter?
- HERRERO No.
- COURT Okay. All right, then we can proceed. We're going to try to get an interpreter.
- BASTOUNES I did try earlier.
- COURT We're trying now, I put in the request.

This transcript was reviewed during an appeal by Hector Herrero of a drug possession conviction and twenty-five year sentence.⁸ The crux of the appeal was since Herrero had a poor understanding of the English language and was not aided by a translator, he was effectively absent from the trial.⁹ While the appeal was not successful, it was not due to Herrero's claims about linguistic ability.¹⁰ Even though the question of whether the transcript reveals that the defendant did not understand English sufficiently to be deemed present at the trial did not

⁸ *Id.* at 242-43.

⁹ *Id.* at 242-45.

¹⁰ *See id.* at 243-44 ("The decision to appoint an interpreter is within the trial court's discretion, and a conviction will be reversed only when an abuse of the court's discretion deprived the defendant of some basic right. . . . Defendant did not ask for an interpreter at any time during jury selection. Nor did defendant raise this issue in his motion for new trial. Consequently, the issue is waived Even if we were to find that this issue was not waived, defendant would still not prevail.").

determine the appeal in this case, it would be beneficial in related circumstances to be able to quantify certainty of interlocutor involvement in the dialogue. Interlocutor involvement is the degree to which dialogue participants are engaged in the conversation; this is more than the frequency and balance of turn-taking, since a virtual monologue with interleaved vocalizations such as “uh-huh” can demonstrate ample turn-taking with little engagement. Notably, in this example, the defendant provided answers to yes-no questions that were coherent as a set. This appears to evidence a greater level of engagement than if those turns were uniformly “yes.” On the other hand, the defendant says little in each question-and-answer to suggest that there was genuine understanding of what the binary responses entailed, except the hedge, “I understand a little bit.” There is no restatement in the defendant’s own words of what was understood. It is precisely a method of quantifying likelihood of understanding that is described and argued appropriate for such forensic contexts in this article. Part III returns to this transcript.

In assessing mutual understanding in conversation, it is important to be clear about what constitutes a null hypothesis and where the burden of proof lies in establishing an alternative hypothesis. Its relevance is illustrated by the fact that the question of whether the evolutionary niche of language is as a cognitive tool for thought or as a cognitive tool for communication remains a topic of debate. This is part of a debate about whether language is a socially evolved construct or a biologically evolved one. It is natural to compare human language and human vision in this context. The eye is a delicate and highly functional product of biological evolution. In contrast, flaws of the linguistic system, including ambiguity at every level of linguistic description, leave language as a poor medium for communication. If one were to design a visual system from scratch, the eye as it is now would likely be a part. If one were trying to evolve a system as well-suited to communication as the eye is for vision, one would strive for telepathy rather than human language.¹¹

¹¹ Cf. Steven Pinker & Paul Bloom, *Natural Language and Natural*

On the other hand, language is a very good system in which to conduct thought, and it serves very well for mental representation and reasoning about a plethora of complex matters. While a speaker may be sometimes vague, typically one who utters an ambiguous sentence has an intended reading in mind. Some topics of potential thought remain notably ineffable, (i.e., thoughts of extreme pain or pleasure or profoundly spatial topics such as geographic directions—precious few people are adept at expressing in words only, without recourse to gesture or maps, how to navigate from one point to another in a city that does not have a grid-based street system), but for the most part, it is difficult to imagine human thought without language.¹²

Given the fundamental flaws of human languages as media for communicating toward mutual understanding, there are strong reasons to view the null hypothesis about human communication in a pessimistic light. In the absence of strong evidence to the contrary, human interaction through dialogue does not reach mutual understanding of the language each other has used in dialogue to describe the world, much less mutual agreement that the world is (or should be) the way that interlocutors understand each other to describe it. A shared understanding of the world may come from common embodiment, the fact that humans share much of their genetic constitution and occupy the same niche in the ecosystem with each other, independently of agreement arising from communication, or from communication nurtured without language used in the process.¹³ I claim that the appropriate null hypotheses about the outcome of language use is not that utterances were interpreted as uttered for all parties to a conversation and agreed in their truth relations to the described world; rather, the null hypothesis pertaining to ordinary dialogue is that communication did not make obvious the existence of disagreement about meanings and the relations between those claims and the world.

Selection, 13 BEHAV. & BRAIN SCI. 707, 784 (1990).

¹² See MICHAEL NEWTON, SAVAGE GIRLS AND WILD BOYS: A HISTORY OF FERAL CHILDREN 20–21 (2002).

¹³ See STEVEN STROGATZ, SYNC: HOW ORDER EMERGES FROM CHAOS IN THE UNIVERSE, NATURE, AND DAILY LIFE 264 (2003).

Literature written about pragmatic theory¹⁴ notes that interlocutors may arrive at working hypotheses by coming to the realization that they have been talking about the same things, understanding each others' comments about those things, and agreeing about the propositional contents put forward on all sides.¹⁵ Much of this literature appeals to processes of *grounding*, which provides a foundation for the speculative conclusions that interlocutors have understood each other.¹⁶ Grounding ultimately is anchored in repetition of words, phrases, and syntactic structures among interacting agents.¹⁷ Detailed conversation analysis has been deployed in legal contexts to emphasize the significant effort necessary to achieve the effect that clients feel understood by their lawyers, for example.¹⁸ Repetitions provide linguistic mechanisms that may be used to develop confidence that conversation has not resulted in misunderstanding.¹⁹ Differential use of repetition according to authority and expertise among interlocutors has been pointed out,²⁰ but in general for all parties in conversation, repeated information is taken to be more securely placed in common ground.²¹

¹⁴ Herbert H. Clark & S.A. Brennan, *Grounding in Communication*, in PERSPECTIVES ON SOCIALLY SHARED COGNITION 127–29 (Lauren B. Resnick et al. eds., 1991). See generally DEBORAH TANNEN, TALKING VOICES: REPETITION, DIALOGUE, AND IMAGERY IN CONVERSATIONAL DISCOURSE (2007).

¹⁵ It is one thing to mutually *understand* that the content of some utterance is the proposition $p \rightarrow q$, and it is another to *agree* that the proposition $p \rightarrow q$ is true.

¹⁶ See generally Paul Vogt & Federico Divina, *Social Symbol Grounding and Language Evolution*, 8 INTERACTION STUD. 31–32 (2007).

¹⁷ Clark & Brennan, *supra* note 14; TANNEN, *supra* note 14.

¹⁸ See, e.g., Linda F. Smith, *Always Judged—Case Study of an Interview Using Conversation Analysis*, 16 CLINICAL L. REV. 423 (2010).

¹⁹ Cf. PATRICK HEALEY, COMMUNICATION AS A SPECIAL CASE OF MISUNDERSTANDING: SEMANTIC COORDINATION IN DIALOGUE (1996).

²⁰ Eve V. Clark & Josie Bernicot, *Repetition as Ratification: How Parents and Children Place Information in Common Ground*, 35 J. CHILD LANGUAGE 349, 364 (2008).

²¹ Clark & Brennan, *supra* note 14; TANNEN, *supra* note 14.

In certain contexts of urgency, grounding mechanisms are part of the ritual of communication designed to avoid miscommunication.²² The rituals of air traffic communication emphasize repetition in order to reduce potential misunderstanding arising from conversation.²³ Repetition of words and phrases has been analyzed as providing a means for interlocutors to increase their involvement in dialogues.²⁴ Accordingly, it follows that enhancing the involvement of aircraft cockpit personnel via repetition increases the chances of a shared understanding of the matter being spoken of by increasing joint immersion in the context at hand. Conscious repetition incrementally eliminates chances that the interlocutors are focused on distinct perspectives on the immediate context.

This article describes and evaluates a method of analysis that can be used to measure engagement in interaction. Where interactions are assessed with respect to these measures, it is possible to quantify certainty that interlocutors have successfully communicated. A growing body of research develops automated and semiautomated methods of measuring synchronization among dialogue participants in terms of such analysis.²⁵ This measure is argued here to correlate with mutual understanding. Some scholars have examined laboratory constructed task-based dialogues in order to correlate effects associated with repetition

²² CUSHING, *supra* note 1, at 40.

²³ *Id.*

²⁴ See TANNEN, *supra* note 14, at 84.

²⁵ See, e.g., Junko Itou & Jun Munemori, *Repetition of Dialogue Atmosphere Using Characters Based on Face-to-Face Dialogue*, 6278 KNOWLEDGE-BASED & INTELLIGENT INFORMATION & ENGINEERING SYS. 527 (2010); Fabian Ramseyer & Wolfgang Tschacher, *Nonverbal Synchrony or Random Coincidence? How to Tell the Difference*, 5967 DEVEL. MULTIMODAL INTERFACES: ACTIVE LISTENING & SYNC. 182 (2010); David Reitter et al., *Computational Modelling of Structural Priming in Dialogue*, PROC. HUM. LANGUAGE TECH. CONF., N. AM. CHAPTER ASS'N FOR COMPUTATIONAL LINGUISTICS ANN. MEETING, 2006, at 121; David Reitter & Johanna D. Moore, *Predicting Success in Dialogue*, PROC. 45TH ANN. MEETING ASS'N FOR COMPUTATIONAL LINGUISTICS, 2007, at 808; Carl Vogel & Lydia Behan, *Measuring Synchrony in Dialog Transcripts*, 7403 COGNITIVE BEHAV. SYS. 73 (2012).

with task-oriented success.²⁶ In the present work, transcripts of dialogues from outside laboratory settings are analyzed. Rather than considering repetition counts up to a point in time from the beginning to an evaluation point in the dialogue, repetitions of tokens as a proportion of total tokens that could have been repeated between an utterance and immediately preceding utterances are considered.²⁷ The level of mutual understanding experienced by the interlocutors is in all cases studied here subjectively assessed, independently through the sources from which the data is drawn. In cases where the method does not support the conclusion that mutual understanding has been achieved, the independent assessments historically provided appear to agree with the conclusions drawn through analysis using the method. The critical cases are those where it is a main issue whether one of the participants understood what was going on, and outside the laboratory environment, it is seldom possible to obtain independent measures of mutual understanding among dialogue participants. Thus, the role of this article is to provide evidence from relatively clear cases that the measurements suggested are valid as a proxy for assessing mutual understanding and to show their efficacy by pointing out the contributions they make in cases that are open to greater debate about the levels of mutual understanding that were likely to have been experienced.

It is not possible to directly measure the actual degree of mutual understanding—neither as a dialogue participant nor as an outside observer. However, the extent to which synchrony and grounding behaviors indicate mutual understanding is the extent to which it may be quantified. If there is no evidence of synchronized engagement, the basis for certainty that there is mutual understanding is undermined. Where even low levels of synchrony are evident, the level of certainty that mutual understanding is in place is correspondingly increased. The method of quantification is to quantify levels of repetition in dialogue. Where repetition differs from chance expectations,

²⁶ See Reitter et al., *supra* note 25, at 122; Reitter & Moore, *supra* note 25, at 809.

²⁷ See Vogel & Behan, *supra* note 25, at 75.

pragmatic inferences may be drawn. The occurrence of such instances has led to the development of a typological theory of functions of repetition.²⁸ The distinction between allo-repetition and self-repetition corresponds to distinctions in the pragmatics of attending to others and attending to oneself. However, this correspondence is not total in that, for example, allo-repetition serves (at least) the purpose of reassuring an initial speaker that a message has been heard but also increasing confidence for the speaker who repeats that the message was at least correctly heard. Self-repetition has a strong relationship with persistence in a dialogue plan. Inversely, allo-repetition avoidance (in its strongest form, a refusal to adopt the language of one's dialogue partners) manifests a focus on the self, and self-repetition avoidance can indicate the absence of an independent dialogue plan. The absence of a significant difference in repetition between actual and randomized dialogues entails a lack of engagement according to that measure and retention of the null hypothesis that mutual understanding was not achieved. This conclusion is based on studies that quantify repetition in mutually engaged conversation²⁹ and communication that leads to success in collaborative tasks.³⁰

This discussion began with an argument that the null hypothesis about linguistic success in ordinary conversation should be that language did not yield mutual understanding; however, in certain legal contexts, the normal burden of proof necessitates even greater caution in assuming that linguistic communication has been successful. The methods discussed here can be used to mitigate risks associated with inappropriately rejecting that null hypothesis. In what follows, a range of dialogues are used to illustrate a theory of repetition in dialogue. The ramifications of the research are as relevant to forensic

²⁸ Note that this is on a different scale of analysis than that of Deleuze of the general epistemological effects of repetition, but is rather restricted to linguistic pragmatics, and as such is a contribution to the framework associated with Tannen. *See generally* GILLES DELEUZE, *DIFFERENCE & REPETITION* (Paul Patton, trans., Athlone Press 2004) (1968); TANNEN, *supra* note 14.

²⁹ *See generally* Vogel & Behan, *supra* note 25, at 73–88.

³⁰ Reitter & Moore, *supra* note 25.

contexts as to safety in air traffic communications: measures of conversational synchrony can be used to mitigate doubt about the extent to which, for example, interrogated individuals have understood the nature of conversations in which they participate, including, for example, the extent of understanding about *Miranda* rights.³¹

The paper is structured as follows. First, a method is described for analyzing interaction in dialogue with respect to interlocutor alignment. The algorithmic core of this method has been employed in the analysis of a range of natural dialogues as recorded in available textual transcripts.³² The essence of the method is the evaluation of the degree to which the various forms of repetition are visible between actual dialogues and a number of randomized reorderings of the dialogue turns. Where precise temporal alignment information is available, overlap and other temporal features of synchronization may also be measured. Lydia Behan and I illustrated how the measures are manifest in natural conversations representative of types: one in which discussion is casual and mutually supportive among participants of equal social standing and another in a crisis situation with a clearly defined leader (an aircraft crash transcript).³³ The work here extends the methods of statistical analysis further and in directions that support forensic deployment of the method in the attribution of interlocutor engagement and understanding of critical legal discourse, for purposes such as police interview, courtroom testimony, and cross-examination. The theory developed here is that where actual repetitions do not exceed random counterparts at all, there is reason to think that the dialogue exemplifies lack of engagement and misunderstanding (or rather, there is no reason to reject the null hypothesis in such a case). Similar assessments are considered with respect to individual participants within the dialogues.

³¹ See generally *Miranda v. Arizona*, 384 U.S. 436, 467 (1966) (holding that an accused person must be made aware of his or her constitutional rights upon arrest, including the right to remain silent).

³² See generally Vogel & Behan, *supra* note 25, at 73–88.

³³ *Id.* at 77–87.

I. METHODS

In all cases, the dialogues analyzed here have already been independently transcribed and are available on the web or in publications cited. The importance of this, given the intention to study repetition across successive dialogue turns, is that the turn-taking structure of dialogue has been determined independently, without evidence of prior reflection on the possibility of this sort of analysis being undertaken. Temporally, the transcripts are partially ordered given that contributions of interlocutors are interleaved; however, temporal overlap analysis is not conducted. Ideally, one would have available not just textual transcripts that indicate the sequence of turns but also the timing of those turns so that temporal overlap of turns can be taken into account. However, as with the dialogues analyzed here, one cannot be guaranteed the availability of timing information.

A decision has to be made with respect to the level of linguistic description at which to consider repetition (tokenization): morphemes, words, part-of-speech (“POS”) labels, concepts, etc. or combinations thereof. The units of representation decided on are types, and their instances are tokens. At this stage, punctuation marks are disregarded. Representation of semantic information that is not directly lexically encoded is not made, since it is not safe to conclude that speakers accept as true all logically valid consequences of their assertions. The text is individuated as words and restricted part-of-speech labeling. POS labeling is only used for personal pronouns to capture the fact that, ordinarily, they are not repeated verbatim but with complementarity, in dialogue that proceeds successfully. Thus, the sole other treatment of the data analyzed here (apart from ignoring punctuation) is to transform dialogues in the form of examples like those numbered below (1) and (3) into those like (2) and (4), respectively; that is, complementary first-person and second-person personal pronouns are replaced with a single item (“IY,” regardless of grammatical number). No deeper parsing is deployed and no other POS labels are used; even third-person pronouns are left intact. Avoiding parsing is desirable to ensure that the methods

are replicable and not dependent on any particular theory of natural language syntax.

- (1) A: Do you understand?
B: I understand.
- (2) A: Is this your address?
B: Yes, this is our address
- (3) A: Do IY understand?
B: IY understand.
- (4) A: Is this IY address?
B: Yes, this is IY address

The tokens counted in the analysis are sequences of words and POS labels of this form; these are known as *n-grams*. The value of *n* varies between one and three. Thus, in the treated dialogue fragment (2) above, B is regarded as repeating two unigrams and one bigram from A's utterance. In (4), B is counted as repeating four unigrams and one bigram. Since they are *sequences*, word order matters, and "this is" does not count as a repetition of "is this," even though there is a natural *syntactic* complementarity between English subject-auxiliary verb inversion in polar questions and the canonical ordering of the subject and verb in answers, just as there is lexical complementarity in pronouns. The equation of the two bigram forms is not made here because of the decision to avoid the need to parse texts. If one were to take syntactic structures into account, then one could consider structural complements as well.

For each transcript processed, the algorithm for data extraction designates a location of memory called a "register" for each speaker. The register, which is initially empty, eventually contains the contents of the most recent contribution of the corresponding speaker.³⁴ A generalization of the method would afford each speaker a vector of registers in order to evaluate repetitions arbitrarily far back in a conversation. In the work described here, the single register for each speaker is initially empty, but it subsequently records the last contribution

³⁴ Register machines or abacus machines with registers form a standard theoretical model of computation. See GEORGE S. BOLOS ET AL., COMPUTABILITY AND LOGIC 45-62 (4th ed. 2002).

made by that speaker. This remains a structural definition, since it is not constrained temporally.

For each utterance, the tokens³⁵ are compared with those recorded in the register for each actor, counting those tokens that are shared between the utterance and its speaker's own register (selfshared; AKA self-repetition) and counting those tokens which are shared with utterances recorded in the registers of each of the other speakers (othershared; AKA allo-repetition). Then the speaker's register is updated to contain the most recently processed utterance. The token counts are conducted for each level of *n-gram*. Repetitions of *n*-grams (again, $1 \leq n \leq 3$) are recorded as counts with respect to the values in the registers as either "SelfShared" or "OtherShared" tokens. In measuring the degree of sharing for a turn u_i , these figures are regarded as proportions of the total number of *n*-grams for each level of *n* that could have been shared, given the length of u_i , between the turn and the immediately preceding turns as recorded in the registers for each actor. In analyzing the three levels of *n* it is useful to think of there actually being two levels, lexical (unigrams) and phrasal (bigrams and trigrams), and therefore this factor is accordingly presented in terms of a derived factor *Nbar* with levels "1" and "2+."

The actual repetition values are then compared with those derived from some number (ten, in each of the experiments here) of randomized reorderings of the turns (AKA contributions). The constituent words within any individual contribution are left intact in their original order; the reordering is of entire turns with respect to other turns into a random partial ordering. This method supports the level of repetition analysis by speaker or aggregated across speakers comparing self-repetition and self-repetition of sequences as manifest in actual dialogues and their turn-randomized counterparts. Forms of repetition are deemed significant in their visibility when the difference between the actual and randomized counterparts is statistically significant.

³⁵ The tokens are in sequences—sequences of words or sequences of POS tags, etc., depending on the choice of tokenization.

Dialogues are reordered by generating new start-times and durations for each utterance as turn indices. The times and durations are selected using random generators based on parameters that depend on the values in the original conversation. Thus, for each utterance u_i a re-indexing \hat{u}_i is constructed. The \hat{u} are sorted on their temporal indices. Analysis of duration of overlaps is enabled by this framework but not conducted here given that the transcripts addressed are not annotated temporally beyond the relative order of turns. Where temporal annotations are available, the analysis of synchronization may have greater depth with the inclusion of consideration of temporal overlap. In the reordered dialogue, counts of allo-shared tokens and self-shared tokens are recorded, just as with actual dialogue. The variables measured and analyzed here are as specified in Table 2. The results allow for the depiction of many contrasting proportions; however, the specific contrasts of interest are whether actual repetition of unigrams and n -grams for larger values of n exceeds the random counterparts for any speaker. Thus, the null hypotheses tested in each dialogue are as in (5) and (6).

$$(5) \text{Randomized.Speaker.1} - \text{Actual.Speaker.1} \geq 0$$

$$(6) \text{Randomized.Speaker.2+} - \text{Actual.Speaker.2+} \geq 0$$

The data is analyzed in each case using a generalized linear model with a binomial error family.³⁶ Adjustments are made for multiple comparisons using directed tests for significance, wherein the null hypothesis essentially is that where DialogType = Randomized repetition will equal or exceed repetition for the corresponding Actual case.³⁷

³⁶ Within R, this is using the following:

```
glm(OSprop ~ DialogType*Speaker*Nbar,family=binomial) and
glm(SSprop ~ DialogType*Speaker*Nbar,family=binomial).
```

³⁷ With the R multcomp package, the following representative constructs are used:

```
fos <- interaction(DialogType,Speaker,Nbar),
mfos <- glm(OSprop ~ fos,family=binomial),
mfos.mc <- glht(mfos,linfct = mcp(fos = "Tukey"),alternative="1").
```

See FRANK BRETZ ET AL., MULTIPLE COMPARISONS USING R (2011). Subsequently, all tests are discarded which do not hold constant Speaker and

Table 2: Variables Analyzed

<i>Variable</i>	<i>Interpretation</i>
DialogType	actual vs. randomized
OtherSpeakers	total number of participants, minus one
n	length of n -grams (1, 2 or 3)
n -bar	n -gram length as a two-level factor ("1" or "2+")
Nbar	n -bar
NGrams	total number of n -grams in a turn
SelfShared	count of tokens from turn shared with own prior contribution
OtherShared	count of tokens from turn shared with prior contribution of other
ss	SelfShared
os	OtherShared
SSrel	SelfShared / NGrams
OSrel	OtherShared / (NGrams * OtherSpeakers)
NonSelfShared	NGrams-SelfShared
NonOtherShared	(NGrams * OtherSpeakers)-OtherShared
nss	NonSelfShared
nos	NonOtherShared
SSprop	SelfShared, NonSelfShared
OSprop	OtherShared, NonOtherShared

II. A CLEAR FAILURE TO COMMUNICATE

The method of analysis advocated here provides an index of synchronized engagement in conversation. This section addresses an example of conversation that has been independently transcribed and analyzed in order to provide a reference point for assessing levels of engagement as a proxy measure of mutual understanding in cases that are selected as more contentious and discussed in Part III. The example is one in which opinion evidently converges on the notion that the conversation does not exemplify mutual understanding.

Nbar and vary solely DialogType.

In a BBC television interview in 1997, Jeremy Paxman interviewed Michael Howard, the former Home Secretary in the UK government.³⁸ The text, prior to treatment as described in Part I, is included in Appendix A. The transcript of the Howard interview is a relatively famous example of an interview in which questions were asked directly, but did not yield responsive answers. If the interlocutors were not both native speakers of English, one might reach the generous conclusion that Howard did not understand the question asked by Paxman. Abstracting over the context, with knowledge of the roles involved—news presenter and politician—one is more likely to infer evasiveness on the part of the politician, since the language used is not manifestly complex. In any case one can conclude from the text of the discussion that the conversation did not instantiate a collaborative flow of information. This conversation is a useful one to show how the proposed measures fare in assessing the level of mutual engagement exemplified. This is demonstrated below.

The mean counts of shared tokens by levels of n are provided in Table 3.

Table 3: Shared token counts in the Paxman-Howard interview

n -bar	<u>OtherShared</u>		<u>SelfShared</u>	
	1	2+	1	2+
Actual	4.71	1.94	3.04	.06
Randomized	4.24	1.58	2.88	.01

The statistical significance of repetitions is assessed by comparing the proportions in token counts that were shared between each utterance and its immediately preceding utterance of the totals that could have been shared in each case. For neither speaker, for neither level of n in the comparisons here is

³⁸ See Interview by Jeremy Paxman with Michael Howard, former U.K. Home Sec’y, *Newsnight* (BBC television broadcast May 13, 1997), available at <http://www.youtube.com/watch?v=Uwlsd8RAoqI>.

the actual repetition (OS,SS) a significantly greater proportion of total tokens that could have been shared (OS+NOS,SS+NSS) for actual dialogue than in the randomized case. The proportions in the comparisons here are depicted as follows: allo-repetition in Figure 1 and self-repetition in Figure 3. The graphs depict the relevant proportions. Of the two sorts of dialogue type, the area occupied by the counts for “Randomized” dialogue is necessarily larger than the area for “Actual” dialogue because there are ten random reorderings of the actual dialogue. Within the sorts of “Other-Sharing” (for either of the two dialogue types) the instances of sharing of items that are shared (“OS”) tends to be much smaller than the number of items that could have been, but were not, “other-shared” (hence, the label, “NOS”). It is apparent that Howard spoke more than Paxman, but the contrast of interest is in the proportions shared and not shared between the actual and randomized conditions for the two individuals. Thus, the mosaic plot³⁹ in Figure 1 does not show any significant difference in allo-repetition for either speaker between the actual and randomized dialogues. The same information, with an additional contrast, is shown in Figure 2: here, the proportion of shared and nonshared unigrams and n -grams, for values of $n > 1$ aggregated, are shown to illustrate the proportions as they depend on the length of expressions, for allo-repetition in this dialog. Recall that the precise statistical tests are used probe ((5) and (6)) for each level of n -bar throughout; however, the graphs which do not separate the levels of n -bar demonstrate the main relationships discussed more clearly. Figure 3, which shows the same proportions for self-repetition, looks different to Figure 1, because Paxman repeated more of his own utterances (“SS”) in relation to his own unrepeated items (“NSS”) than Howard repeated of his own utterances. However, for neither Paxman nor Howard is the difference significantly greater for the actual dialogue than the randomized counterparts.

³⁹ David Meyer et al., *The Strucplot Framework: Visualizing Multi-Way Contingency Tables with VCD*, J. STAT. SOFTWARE, Oct. 2006, at 1 (“A mosaic plot is basically an area-proportional visualization of (typically, observed) frequencies, composed of tiles (corresponding to the cells) created by recursive vertical and horizontal splits of a rectangle.”) (citations omitted).

Table 4: 1997 Paxman-Howard interview: counts of shared and nonshared tokens by dialogue-type, speaker, and n-bar

Speaker	DialogType	OtherSharing				SelfSharing			
		Actual		Randomized		Actual		Randomized	
		Nbar	OSNOS OS	NOS	SS NSS	SS	NSS		
Howard	1	43	154	435	1435	46	151	423	1547
	2+	1	357	4	3576	12	346	125	3455
Paxman	1	30	91	256	954	67	54	594	616
	2+	2	204	3	2057	81	125	631	1429

Figure 1: Allo-repetition by Paxman & Howard: shared vs. nonshared tokens by speaker in actual and randomized dialogue

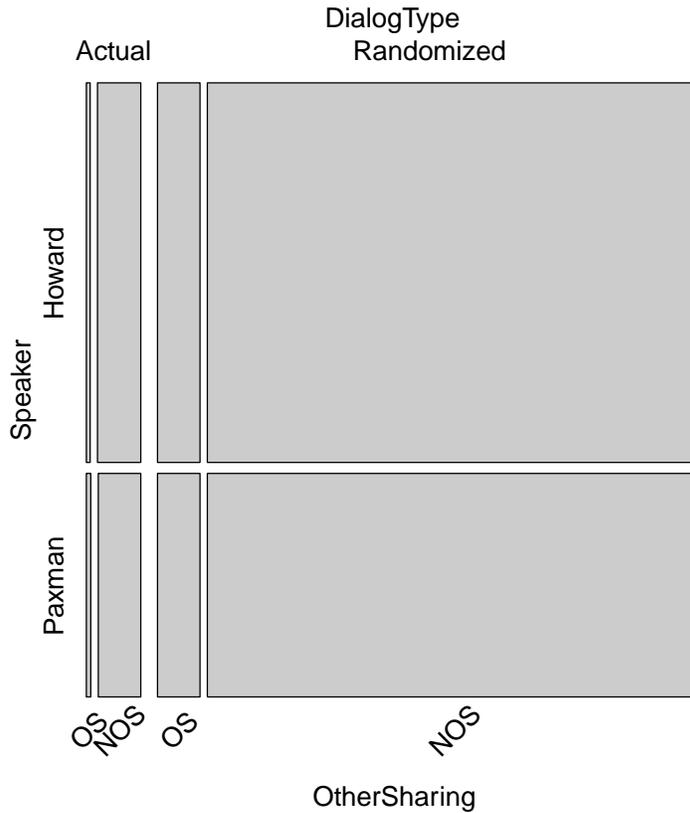


Figure 2: Allo-repetition by Paxman & Howard: shared vs. nonshared tokens by speaker and n-bar in actual and randomized dialogue

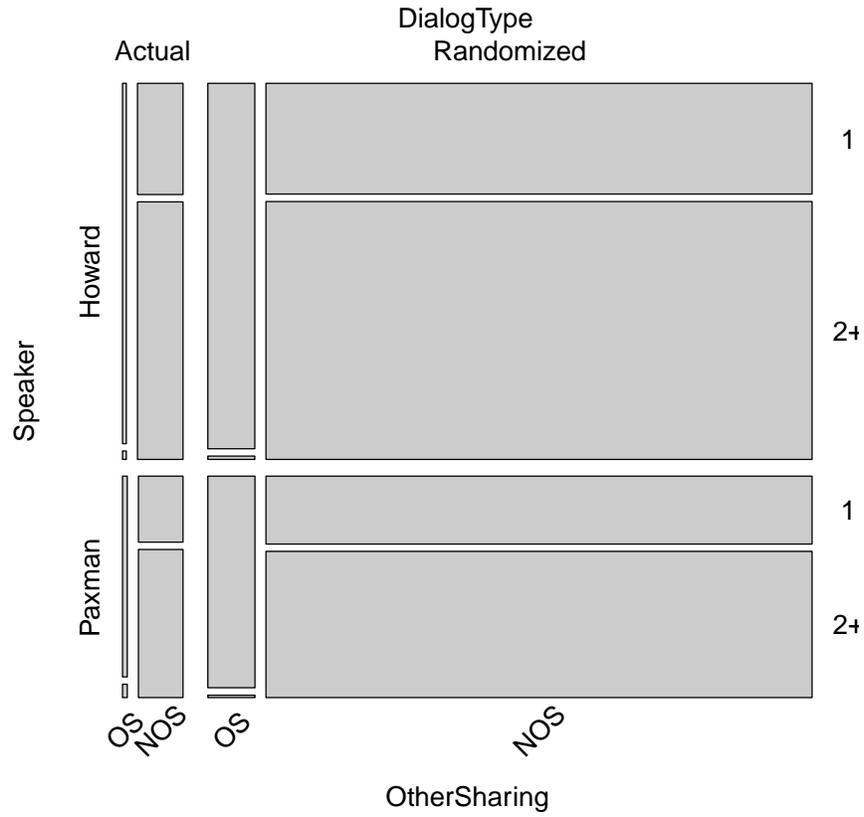
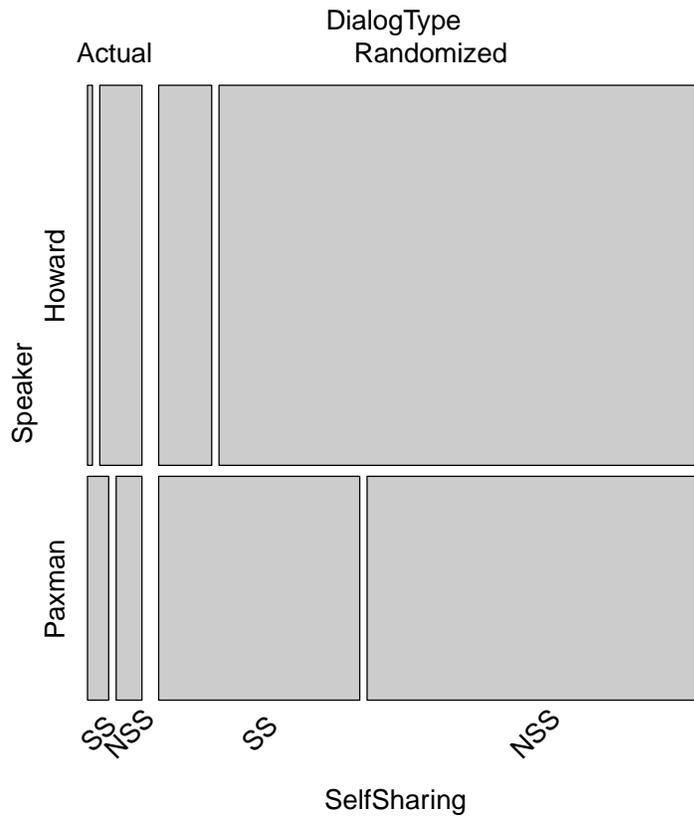


Figure 3: Self-repetition by Paxman & Howard: shared vs. nonshared tokens by speaker in actual and randomized dialogue



The lack of significance in the difference between the actual and randomized dialogues in the direction taken as of interest here suggests that the dialogues might as well have had the turns randomly reordered to obtain the same overall effect of engagement. Certainly, it is clear that the repetitions of the Paxman questions makes self-repetition high for Paxman but not substantially different in the ten randomized reorderings. Similarly, Howard's repetition of terms used by Paxman is relatively impervious to reordering. The statistical effects reveal that at the level of textual content, there is little engagement exhibited. Accordingly, this leaves open that an analysis including gesture, timing of utterances, or overall energy measurable in the scene during the flow of the dialogue could still detect involvement and engagement at that level,⁴⁰ just as speakers of mutually unintelligible languages may interact with engagement but without full understanding. The lack of significance in the contrasts of interest here implies that, although it seems that willful avoidance might be at issue, one cannot say for certain that Paxman and Howard reached an understanding of each other.

It is interesting to note that it is relatively difficult to find transcripts of naturally, publicly occurring dialogues in which third party observers conclude that the interlocutors do not understand each other. Fabrications of such are the stuff of comedy, such as the "Who's on first" routine of Abbott and Costello. The Paxman-Howard example was selected as a relatively famous example of failure to communicate. With more successful dialogues, it is generally not true that in all cases all parties will support measurements according to the methods used here in which the actual dialogue differs in the hypothesized direction from the ten turn-randomized alternatives. Where there is significant self-repetition but no significant allo-repetition for any of the participants in a dialogue, it would seem that there is evidence of persistence but not of linguistic engagement. Conversely, where allo-repetition effects are significant, but not self-repetition, there is evidence of understanding and engagement. Where neither of the effects is visible for any of

⁴⁰ Ramseyer & Tschacher, *supra* note 25.

the participants, the conversation is difficult to classify as successful. It is not safe, in such a case, to make judgments about the level of mutual understanding achieved among the participating parties.

III. LEVELS OF MUTUAL UNDERSTANDING IN SAMPLE CASES

The next two sections present analyses of transcripts which represent the sorts of situations in which one might expect the methods described here to have benefit. In the present context, they constitute evaluations of the method to the extent that the conclusions supported by the method are in agreement with independent historical assessments of the conversations.

A. *Case Study 1: People v. Herrero*⁴¹

The data the first case study relies upon has been described in the introduction.⁴² The data is treated and processed according to the methods specified, with the actual dialogue giving rise to ten counterparts in which the turns have been randomly reordered with respect to each other.

1. *Results*

As before, the proportions of allo-repetition⁴³ and self-repetition⁴⁴ are analyzed. Notably, Herrero has higher levels of both OS and SS in the randomized versions of the dialogue than in the actual dialogue. Furthermore, recall that the testing conducted is directional and the null hypothesis is that random repetition will exceed or equal actual repetition in the proportion measures ((5) and (6)).

⁴¹ *People v. Herrero*, 756 N.E.2d 234, 242–43 (Ill. App. Ct. 2001).

⁴² *See supra* Introduction.

⁴³ *See infra* Figure 4.

⁴⁴ *See infra* Figure 5.

Figure 4: Allo-repetition in the Herrero case: shared vs. nonshared tokens by speaker in actual and randomized dialogue

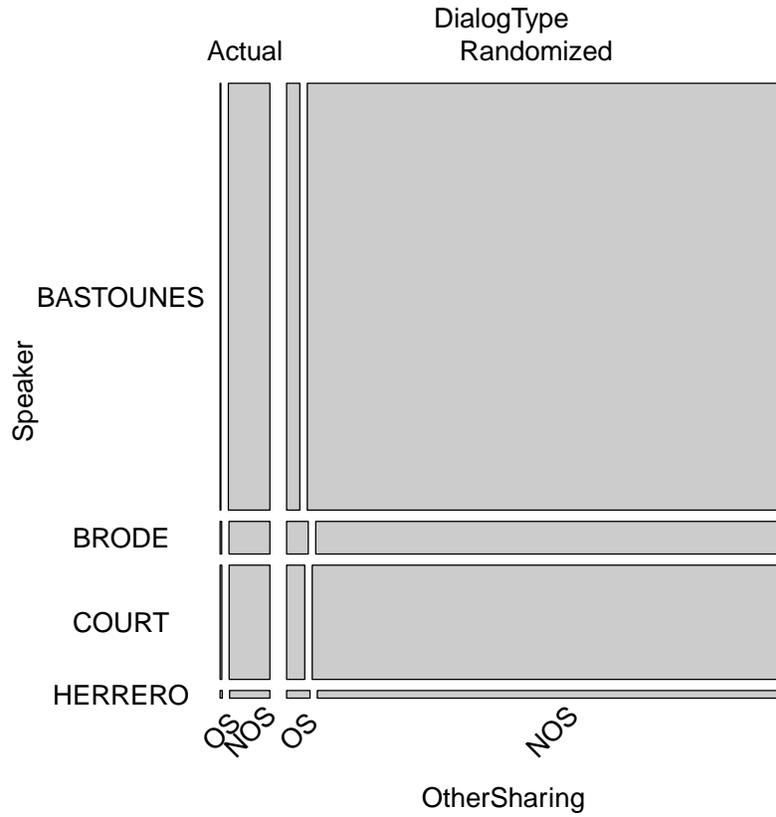


Figure 5: Self-repetition in the Herrero case: shared vs. nonshared tokens by speaker in actual and randomized dialogue

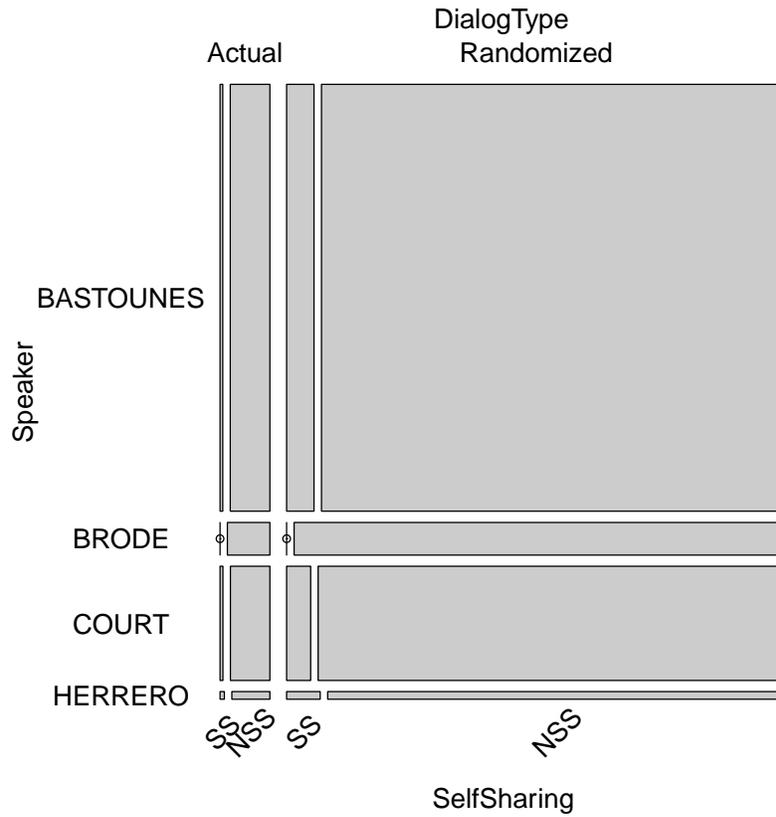


Table 5: Herrero case: counts of shared and nonshared tokens by dialogue-type, speaker, and n-bar

Speaker	DialogType Nbar	<u>Other Sharing</u>				<u>Self Sharing</u>			
		Actual		Randomized		Actual		Randomized	
		OS	NOS	OS	NOS	SS	NSS	SS	NSS
BASTOUNES1	1	63	613	737	6073	52	175	490	1780
	2+	7	2450	105	24465	16	803	92	8098
BRODE	1	7	47	92	448	0	18	0	180
	2+	2	184	14	1846	0	62	0	620
COURT	1	34	164	288	1692	17	49	124	536
	2+	2	640	223	6398	3	211	13	2127
HERRERO	1	3	24	27	243	2	7	13	77
	2+	0	30	0	300	0	10	0	100

Recall that the null hypothesis asserts that there is no more repetition in actual dialogue than in randomized counterparts. The contrasts between actual and randomized dialogues are not significant enough to allow rejection of the null hypothesis for any participant nor for any level of n -bar, whether for self-repetition or allo-repetition.

2. Discussion

The results of applying the method suggest that the null hypothesis must be retained: the dialogue does not present sufficient repetition of words or phrases to suggest that the interlocutors have engaged sufficiently to achieve mutual understanding. Despite the fact that the answers to the questions posed provided by Herrero are rational and mutually consistent, particularly given that the individual is a nonnative speaker of English and given that the answers are all one-word responses to polar interrogatives, there simply is not sufficient evidence here and on these measures to support the claim that Herrero understood the proceedings. In the actual legal case, as discussed in the introduction, the final decision did not hinge on the answer to the question of whether the defendant understood

the proceedings. The dialogue was produced to the Illinois Appellate Court, First District, with an implicit argument that the responses of the defendant were contextually appropriate, but with a null hypothesis about whether this is evidence of mutual understanding that is at odds with the arguments given in the introduction.⁴⁵

*B. Case Study 2: State v. Cunningham*⁴⁶

The Supreme Court of North Carolina heard an appeal in which the defendant sought a new trial because of faulty jury selection. The transcript of jury selection was considered to assess whether a particular potential jury member, Carnes, whose exclusion from the jury was not permitted, had been predisposed towards a particular verdict or understood the notion of presumed innocence. Defense counsel Murphy tried to explain the notion of presumed innocence and attempted to ascertain whether Carnes understood. Occasionally, Wolfe, for the prosecution, and the Court intervened. As an independent reader of the transcript of the relevant jury selection process, it is easy to form the opinion that Carnes did not understand what it meant to accept the concept of presumed innocence.⁴⁷

1. Results

Table 6 shows the distribution of token counts across the categories studied. Figure 6 depicts the relative proportions of shared vs. nonshared tokens in the case of allo-repetition. None of the contrasts of interest are statistically significant for allo-repetition (i.e., actual vs. randomized tokens produced by each speaker for each level of *n*-bar, repeating tokens from the last turns of all of the other speakers). Figure 7 shows the same

⁴⁵ *Herrero*, 756 N.E.2d 234.

⁴⁶ *State v. Cunningham*, 474 S.E.2d 772 (N.C. 1996). The transcript is included in App. B. The data used here is taken from the public records of the State of North Carolina. *State v. Cunningham No. 232A91*, IBIBLIO, <http://www.ibiblio.org/pub/docs/nc-supreme-court/jun0493/cunningham.asc> (last visited Apr. 17, 2013).

⁴⁷ See *State v. Cunningham No. 232A91*, *supra* note 46.

proportions for self-repetition. Self-repetition effects are significant for Murphy (the defense attorney questioning the potential juror) and Carnes (the juror being questioned) in that actual repetition exceeds repetition in randomized dialogues, for both unigrams and 2+ grams (adjusted $p \leq .01$). The effect also exists for the court with respect to unigrams (adjusted $p \leq .01$). Actual repetition of 2+ grams is not significantly in excess of randomized repetition for the court, and neither level of n -gram yields significant actual self-repetition for Wolfe (the prosecutor).

Figure 6: Allo-repetition in the Cunningham case: shared vs. nonshared tokens by speaker in actual and randomized dialogue



Figure 7: Self-repetition in the Cunningham case: shared vs. nonshared tokens by speaker in actual and randomized dialogue

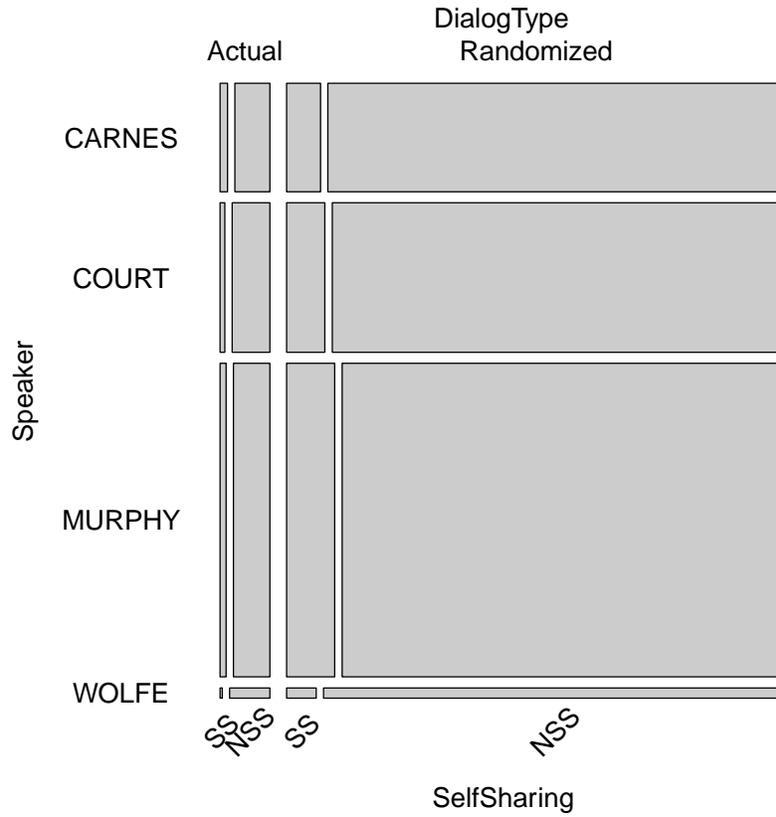


Table 6: State v. Cunningham: counts of shared and nonshared tokens by dialog-type, speaker, and n-bar

Speaker	DialogType Nbar	Other Sharing				Self Sharing			
		Actual OS	Actual NOS	Randomized OS	Randomized NOS	Actual SS	Actual NSS	Randomized SS	Randomized NSS
CARNES	1	159	933	1627	9293	110	254	558	3082
	2+	30	1827	285	18285	65	554	119	6071
COURT	1	175	1238	2029	12101	110	361	761	3949
	2+	32	2623	315	26235	42	843	295	8555
MURPHY	1	261	2679	4023	25377	324	656	2462	7338
	2+	23	5575	451	55529	80	1786	316	18344
WOLFE	1	11	100	163	947	4	33	46	324
	2+	1	167	14	1666	1	55	10	550

2. Discussion

Applying the method proposed to the data examined here supports the conclusion that the prosecutor and potential juror persisted in their respective communication strategies. The fact that allo-repetition effects that distinguish the actual dialogue from its ten randomized counterparts do not appear strengthens the argument that mutual understanding did not emerge during the interaction. This analysis coincides with the determination on appeal that the jury member did not unambiguously demonstrate understanding of the concept of presumed innocence.

IV. CONCLUSIONS

While the current work is in the spirit of the traditions in statistical methods for authorship attribution,⁴⁸ it is focused on

⁴⁸ See generally GEORGE U. YULE, *THE STATISTICAL STUDY OF LITERARY VOCABULARY* (1944) (examining word-distribution from different portions of author's works throughout history); Harald Baayen et al., *Outside the Cave of Shadows: Using Syntactic Annotation to Enhance the Authorship Attribution*, 11 *LITERARY & LINGUISTIC COMPUTING* 121, 121-32 (1996) (reporting an experiment in which statistical measures and methods previously applied to words and their frequencies of use are applied to rewrite rules);

the attribution of mutual understanding and engagement in dialogue. The distinguishing features sought are not linguistic (words and phrases) in themselves but patterns of their use. Hence, the method can be considered to provide a tool for quantifying pragmatics. Other algorithmic means have been used to assess levels of understanding and awareness in dialogue. Apart from the explicit use of lexicalized feedback tags (e.g., “eh?”, “I see”), information contained in prosody has also been studied.⁴⁹ In analysis of other dialogues, the methods underlying

Carole E. Chaski, *Empirical Evaluations of Language-Based Author Identification Techniques*, 8 INT’L J. SPEECH LANGUAGE & L. 1, 1–65 (2001) (testing language-based author identification techniques based on syntactic analysis, syntactically classified punctuation, sentential complexity, vocabulary richness, readability, content analysis, and errors); Carole E. Chaski, *Who’s at the Keyboard? Authorship Attribution in Digital Evidence Investigations*, 4 INT’L J. DIGITAL EVIDENCE, Spring 2005, at 1, 1–13 (applying computational, stylometric authorship attribution methods to crimes involving digital evidence); Jack Grieve, *Quantitative Authorship Attribution: An Evaluation of Techniques*, 22 LITERARY & LINGUISTIC COMPUTING 251, 251–70 (2007) (comparing thirty-nine different types of textual measurements commonly used in authorship attribution studies to determine which measurements are the best indicators); David I. Holmes, *Authorship Attribution*, 28 COMPUTERS & HUMAN. 87, 87–106 (1994) (quantifying literary style and looks at several variables to find the stylistic “fingerprints” of a writer); Kim Luyckx & Walter Daelemans, *Shallow Text Analysis and Machine Learning for Authorship Attribution*, PROC. FIFTEENTH MEETING COMPUTATION LINGUISTICS IN NETH., 2005, at 149–60, *available at* <http://lotos.library.uu.nl/publish/articles/000139/bookpart.pdf> (reporting on the use of syntax-based features as possible predictors for an author’s style and token-based features that are predictive to author style); Harold Somers & Fiona J. Tweedie, *Authorship Attribution and Pastiche*, 37 COMPUTERS & HUMAN. 407, 407–29 (2003) (testing whether authorship attribution techniques can distinguish between a deliberate imitation and its model); Carl Vogel & Gerald Lynch, *Computational Stylometry: Who’s in a Play?*, in VERBAL AND NONVERBAL FEATURES OF HUMAN-HUMAN AND HUMAN-MACHINE INTERACTION 1, 169–86 (2008) (applying automatic text classification techniques to quantifying strength of characterization within plays); George U. Yule, *On Some Properties of Normal Distributions, Univariate and Bivariate, Based on Sums of Squares of Frequencies*, 30 BIOMETRIKA 363, 363–90 (1938) (evaluating univariate and bivariate distributions and squaring every ordinate).

⁴⁹ See generally Jens Edlund et al., *The Effects of Prosodic Features on the Interpretation of Clarification Ellipses*, 2005 PROC. INTERSPEECH 2389

the current work yield significant differences between actual and randomized repetitions, which support the inference that significant engagement has happened, giving more certainty to attributions of mutual understanding,⁵⁰ thereby supporting the face validity of the methods. Further, recall that face validity is found elsewhere, in the correlation between repetition levels and task success in task-based dialogue.⁵¹

The capacity to assess and attain mutual understanding from a position external to a dialogue is important outside forensic contexts. The capacity to reliably and objectively make such attributions is also relevant in clinical diagnostics. Moreover, even if repetition behaviors do not figure explicitly into Diagnostic and Statistical Manual diagnoses of schizophrenia, analysis of repetition figures into current concepts of schizophrenia,⁵² including attention to whether such analyses can lead to inappropriate diagnoses.⁵³ In preliminary study in this

(describing an experiment in which subjects listened to short dialogue fragments and judged what was actually intended by the elliptical clarification, based on prosodic features); Jens Edlund et al., *User Responses to Prosodic Variation in Fragmentary Grounding Utterances in Dialog*, PROC. INTERSPEECH 2006, at 2002 (testing whether subjects change their behavior to different fragmentary grounding utterances in a human-computer dialogue setting).

⁵⁰ Vogel & Behan, *supra* note 25.

⁵¹ See Reitter & Moore, *supra* note 25, at 808–15.

⁵² See generally Michael Covington et al., *Schizophrenia and the Structure of Language: The Linguist's View*, 77 SCHIZOPHRENIA RES. 85, 85–98 (2005) (evaluating schizophrenic language impairments to see how schizophrenia affects phonology, syntax and semantics); Gina R. Kuperberg, *Language in Schizophrenia Part 1: An Introduction*, 4 LANGUAGE & LINGUISTICS COMPASS 576, 576–89 (2010) (discussing language output in schizophrenia and the theory that language dysfunction in schizophrenia arises from abnormalities in semantic memory and/or working memory and executive function); Gina R. Kuperberg, *Language in Schizophrenia Part 2: What Can Psycholinguistics Bring to the Study of Schizophrenia . . . and Vice Versa?*, 4 LANGUAGE & LINGUISTICS COMPASS 576, 590–604 (2010) (applying online psycholinguistic methods to schizophrenic language).

⁵³ See Susan Trumbetta et al., *Language-Related Symptoms in Persons with Schizophrenia and How Deaf Persons May Manifest These Symptoms*, 1 SIGN LANGUAGE STUDIES 214, 228–53 (2001).

area,⁵⁴ transcripts previously analyzed by Steuber⁵⁵ in the identification of linguistic features that discriminate transcripts of interviews with individuals diagnosed with either schizophrenia or depression have been considered, finding no significant level of repetitions of others, but individual persistence via self-repetition. The potential use of the methods proposed in the suite of tools for diagnosis of syndromes with distinctive accompanying effects on conversational linguistic abilities is an area ripe for deeper exploration using the methods demonstrated here.

This article has introduced a method of interaction analysis based on repetition analysis that is distinct in analytical details from other analytical methods in the literature. The use of the methods has been demonstrated by analyzing transcripts that are freely available and with respect to which it is possible to draw upon independent assessments of the degree to which the transcribed conversations demonstrate engaged interaction and mutual understanding. Allo-repetition effects are taken to be those where the repetitions of tokens by an individual of dialogue partners immediately preceding contributions, summed over the conversation, significantly exceed in actual conversation the same measurement averaged over turn-randomized treatments of the conversation. Self-repetition effects are those where in the cumulative counts of repetitions of a speaker's immediately prior contribution significantly exceed for actual conversation the averaged accumulated counts for randomized counterparts. Self-repetition effects are taken to be indicative of speaker persistence with dialogue plans. Allo-repetition effects are taken to be signals of mutual engagement in dialog, and the conversations where these effects appear accord with independent intuition about the level of engagement and mutual understanding (distinct from mutual agreement) achieved within

⁵⁴ Carl Vogel, *Quantifying Interaction Synchrony as Evidence of Mutual Understanding*, 49 CORTEX (forthcoming 2013).

⁵⁵ Lucas C. Steuber, *Disordered Thought, Disordered Language: A Corpus-Based Description of the Speech of Individuals Undergoing Treatment for Schizophrenia* (2011) (unpublished M.A. thesis, Portland State Univ.), http://dr.archives.pdx.edu/xmlui/bitstream/handle/psu/7087/Steuber_psu_0180_E_10321.pdf?sequence=1.

the conversation. Thus, the method is put forward for further exploration in contexts that require quantified analysis of attributions of mutual understanding in linguistic interaction. The argument made here is that where actual allo-repetition in a dialogue is not significantly in excess of its counterpart measure in randomized versions of dialogue, the level of engagement is insufficient to make confident attributions of mutual understanding.

Appendix

A. Transcript of Interview with Michael Howard⁵⁶

Paxman Did you threaten to overrule him?
 Howard I was not entitled to instruct Derek Lewis and I did not instruct him.
 Paxman Did you threaten to overrule him?
 Howard The truth of the matter is that Mr. Marriot was not suspended
 Paxman Did you threaten to overrule him?
 Howard I did not overrule Derek Lewis
 Paxman Did you threaten to overrule him?
 Howard I took advice on what I could or could not do
 Paxman Did you threaten to overrule him?
 Howard and acted scrupulously in accordance with that advice. I did not overrule Derek Lewis
 Paxman Did you threaten to overrule him?
 Howard Mr. Marriot would not suspend him
 Paxman Did you threaten to overrule him?
 Howard I have accounted for my decision to dismiss Derek Lewis
 Paxman Did you threaten to overrule him?
 Howard in great detail before the House of Commons
 Paxman I note that you're not answering the question whether you threatened to overrule him.
 Howard Well, the important aspect of this which it's very clear to bear in mind
 Paxman I'm sorry, I'm going to be frightfully rude but—I'm sorry—it's a straight yes-or-no question and a straight yes-or-no answer did you threaten to overrule him?
 Howard I discussed the matter with Derek Lewis. I gave him the benefit of my opinion. I gave him the benefit of my opinion in strong language, but I did not instruct him because I was not, er, entitled to instruct him. I was entitled to express my opinion and that is what I did.

⁵⁶ Yuri Prime, *Newsnight*, EVERYTHING2, <http://everything2.com/title/Newsnight> (last visited Apr. 17, 2013).

Paxman With respect, that is not answering the question of whether you threatened to overrule him.

Howard It's dealing with the relevant point which was what I was entitled to do and what I was not entitled to do, and I have dealt with this in detail before the House of Commons and before the select committee.

Paxman But with respect you haven't answered the question of whether you threatened to overrule him.

Howard Well, you see, the question is. . . .

*B. Transcript of Contested Jury Member Selection in State v. Cunningham.*⁵⁷

MURPHY Do you understand, Ms. Carnes, that we have at law what is called the presumption of innocence, that is, a person who is charged with a criminal offense is presumed to be innocent until and unless the State can prove that person's guilt beyond a reasonable doubt?

CARNES Yes, sir.

MURPHY You understand that, don't you?

CARNES Yes, sir, I do.

MURPHY And, of course, you understand that the charge in this particular case is first-degree murder. It involves the shooting of a police officer. Do you understand that?

CARNES Yes, I do.

MURPHY And one of the things that you will be called upon to do is to apply the principles that we were talking about to this particular case if you sit as a juror.

CARNES Yes, I do.

MURPHY Now, it is one thing, of course, to say that you can do something and it may be entirely different.

CARNES Yes.

MURPHY That is, that you actually be able to do that, and that is really what I want you to search yourself about. I

⁵⁷ State v. Cunningham No. 232A91, *supra* note 46.

want you to think about that. You seem to be one who holds your opinion strong, and that's fine. Given that you have such a strong feeling about the death penalty in your statement that if a person takes another life, they should be put to death, given that Mr. Cunningham is charged with first-degree murder, as you sit there today, can you honestly say to yourself, not to me necessarily but to yourself, that you are able to presume Mr. Cunningham innocent?

CARNES Until he is proven guilty.

MURPHY Do you expect that to happen?

WOLFE Object.

COURT Sustained.

CARNES I don't know.

COURT Don't answer the question when I sustain it.

MURPHY I understand that if he is proven guilty of first-degree murder, then that would remove the presumption of innocence, but that is really not what I am asking you. Okay? What I am really asking you at this point is can you honestly, as he sits there right now, and as you sit in that seat right now, and nobody knows this any better than you, I'm just asking, can you honestly presume him to be innocent?

CARNES Yes, because I don't know what happened.

MURPHY Now, part and parcel of the principle of the presumption of innocence is the defendant's right not to testify, not to present any evidence if he doesn't want to, because he doesn't have that burden. The State has the entire burden of proof in a criminal case to satisfy you beyond a reasonable doubt of a person's guilt, if they can do that. Okay? Now, would it present a problem for you in returning a verdict of not guilty if the State fails to prove to you beyond a reasonable doubt the defendant's guilt and Mr. Cunningham didn't testify?

CARNES I'm not sure I follow that.

MURPHY Okay. If Mr. Cunningham doesn't testify in this

case, in your mind does that make the State's job any more difficult or easier?

CARNES I would think it would be more difficult.

MURPHY If he does not testify?

CARNES Yes, because they have to prove him innocent or guilty. I would think that he would have to testify, or need to.

MURPHY Okay. You understand that the State only has to prove guilt. They don't have to prove innocence.

CARNES Yes.

MURPHY And is it your expectation or is it your thinking now that we would have to prove that Mr. Cunningham is innocent?

CARNES Do I think you would have to prove it?

MURPHY Yes.

CARNES Yes, I thought that is what you would be trying to do.

MURPHY Trying to prove that he's innocent?

CARNES Yes.

MURPHY Do you understand that the burden of proof is on the State?

CARNES Yes.

MURPHY Not us?

CARNES Yes.

MURPHY You would still expect him, or us, to prove Mr. Cunningham is innocent. Correct?

CARNES Yes.

WOLFE I would ask for a clarification on the law on that, your Honor.

COURT Ms. Carnes—

CARNES He's getting me very confused.

COURT Okay. Let me explain to you. I think I told you that Mr. Cunningham has entered a plea of not guilty.

CARNES Yes.

COURT And under the law of North Carolina, the fact that he has been charged with a crime is not evidence of his guilt. He is not required to prove his innocence; he is presumed to be innocent.

CARNES Okay.

COURT The State of North Carolina has the burden of proof, and that burden is to prove each element of the offense of which he is charged beyond a reasonable doubt. Now, the law also says that Mr. Cunningham does not have to testify in his own behalf. He doesn't have to call any witnesses or present any other form of evidence, and that you cannot hold that against him. Do you understand that?

CARNES Yes, ma'am.

COURT Can you follow that law?

CARNES Yes, ma'am.

COURT Mr. Murphy?

MURPHY All right. Now, that's what I'm asking you, Ms. Carnes. The judge told you what the law is, and I think the district attorney also said the same thing to you. I thought I had explained that. I thought I said that. Now, the question is your ability to follow that law.

CARNES Yes.

MURPHY And that's what I'm asking you. That given your understanding at this point—and I trust that that is clear – is it your feeling that Mr.—we at this table would have to prove to you that Mr. Cunningham is innocent of this offense?

CARNES Yes.

MURPHY We offer her for cause.

WOLFE Object, Your Honor.

COURT Did you understand the explanation?

CARNES Yes, ma'am.

COURT And in light of my explanation that he is presumed to be innocent and is not required to prove his innocence, you would still require him to testify or to prove his innocence?

CARNES Right now he is innocent, or he is innocent until proven guilty. I understand that. But you are saying I need to—I'm sorry, I'm not sure.

COURT You need to slow down just a little bit.

CARNES He is innocent until proven guilty. I understand that, until he is proven guilty, before we can say he is

guilty. That, I understand.

COURT Which part is it that you don't understand?

CARNES Well, I thought I understood everything.

COURT Well, I told you that he is not required to prove his innocence.

CARNES Then I guess that means his attorney will have to prove that he is not guilty? He doesn't have to prove his innocence then—is that what you're saying—since he's innocent until proven guilty?

COURT Let me start over. I told you that the fact that he has been charged with an offense is not evidence of his guilt. You can't consider it as evidence of his guilt. I told you also that he is presumed to be innocent and is not required to prove his innocence. The State of North Carolina, represented by Mr. Wolfe and Ms. Brown, has the burden of proof. That burden is to prove each element of the offense with which Mr. Cunningham is charged beyond a reasonable doubt. The State has to carry that burden of proof and convince all twelve jurors beyond a reasonable doubt of each element of the offense before the jury may return a verdict of guilty. Mr. Cunningham is presumed to be innocent, and that presumption stays with him throughout the course of the trial unless the jury finds after they go into the jury room to deliberate that the State has carried its burden of proof. The law also says that Mr. Cunningham does not have to testify. He does not have to call any witnesses on his behalf or present any evidence. He is not required to prove his innocence. And that you, as a juror, cannot hold that against him. Do you understand that?

CARNES Yes.

COURT Were you confused?

CARNES Yes.

COURT I'm going to deny the challenge for cause at this point.

MURPHY Okay. Ms. Carnes, it is not my purpose to try to confuse you. That's why I want you to stop me

when we go along. If you don't understand anything that I have said, or if you need further clarification, stop me and we will ask the judge to do that because we don't want a confused juror. We want a juror who is clear with what they have to do. Okay?

CARNES Okay.

MURPHY Now, I do, however, want to pursue that with you just a little bit because I want to know how you feel about the matter and not just telling me things because you think that's what I want to hear. Okay? Because it's not what I want to hear; it's how you honestly feel about things. And what I want you to tell me is that if you would require the defendant to prove his innocence to you.

CARNES No.

MURPHY You would be satisfied then just to hear from the State and rely, if necessary, just on what the State presents to you on the guilt or innocence phase before you would return a verdict. Is that correct?

WOLFE Object.

MURPHY Well, the State has the burden.

WOLFE That is an improper statement of the law.

COURT Sustained as to form.

MURPHY I will rephrase the question. Can you require the State to prove to you, if they can, Mr. Cunningham's guilt beyond a reasonable doubt?

CARNES Well, if I understand what they are saying, they have to prove he is guilty and not require his innocence to be proven. He doesn't have to prove his innocence, I guess, is what I'm trying to say.

MURPHY And would you accept that? I mean—

CARNES Yes, if he doesn't want to prove his innocence, I would have to accept that.

MURPHY Okay.

WOLFE May we approach the bench just a minute, Your Honor?

COURT Yes.

MURPHY I guess I'm a little bit confused myself at this point, Ms. Carnes. Let's see if we can understand each

other. Okay? You had indicated something to the effect that if we didn't want to prove his innocence, that you would accept what the State offered?

CARNES I understand that he is innocent right now until proven guilty. So if they prove him guilty, I would accept the fact that he is guilty, if they prove him to be guilty.

MURPHY Okay. I guess I didn't understand what you meant when you said if we didn't want to prove his innocence.

CARNES And then I said until they prove him guilty. When they prove him guilty, then he is guilty, when they prove he is guilty.

MURPHY Okay. When you say when they prove him guilty, what do you mean?

CARNES When they prove that he did it, when they come up with all of the evidence that he did it.

MURPHY I suppose I'm having some problems with that. It sounds like you expect them to do that.

WOLFE Object.

CARNES No, I don't. I said—

COURT Sustained.

CARNES Well, I should have said if they do.

COURT Ms. Carnes, let me say when there is an objection, you need to stop talking.