Complex Questions Asked by Defense Lawyers But Not Prosecutors Predicts Convictions in Child Abuse Trials

Angela D. Evans,
Institute of Child Study, University of Toronto, 45 Walmer Road, Toronto, ON, Canada M5R 2X2

Kang Lee, and
Institute of Child Study, University of Toronto, 45 Walmer Road, Toronto, ON, Canada M5R 2X2 kang.lee@utoronto.ca

Thomas D. Lyon
USC Gould School of Law, University of Southern California, 699 Exposition Blvd., Los Angeles, CA 90089-0071, USA

Abstract

Attorneys' language has been found to influence the accuracy of a child's testimony, with defense attorneys asking more complex questions than the prosecution (Zajac & Hayne, J. Exp Psychol Appl 9:187–195, 2003; Zajac et al. Psychiatr Psychol Law, 10:199–209, 2003). These complex questions may be used as a strategy to influence the jury's perceived accuracy of child witnesses. However, we currently do not know whether the complexity of attorney's questions predict the trial outcome. The present study assesses whether the complexity of questions is related to the trial outcome in 46 child sexual abuse court transcripts using an automated linguistic analysis. Based on the complexity of defense attorney's questions, the trial verdict was accurately predicted 82.6% of the time. Contrary to our prediction, more complex questions asked by the defense were associated with convictions, not acquittals.

Keywords

Linguistic; Testimony; Lawyers; Child

Each year approximately 100,000 children testify in the United States of America (Bruck, Ceci, & Hembrooke, 1998). Within criminal court, children most often testify about sexual abuse (Goodman, Quas, Bulkley, & Shapiro, 1999). Unfortunately, sexual abuse allegations are often difficult to prove; eyewitnesses are uncommon and physical evidence, when it exists, rarely points to a specific perpetrator. Thus the child witnesses' testimony is an influential factor for jury members when determining the trial outcome.

While testifying, child witnesses are often faced with answering complex and confusing questions. Language used by attorneys has been found to be developmentally inappropriate (Brennan & Brennan, 1988; Cashmore & DeHaas, 1992; Flin, Stevenson, & Davies, 1989;

---

[1]A preliminary logistic regression was performed including the prosecutions' questions. The verdict was the predicted variable with Age entered on the first step, followed by mean wordiness scores for the defense and prosecution and finally the mean defense and prosecution complexity scores for each of the six child response categories. None of the prosecution's complexity scores significantly predicted the trial outcome, all p's < .05.

Goodman, Taub, Jones, & England, 1992; Peters & Nunez, 1999). It tends to contain legalistic terminology and complex sentence forms, such as double negatives and tag questions. Legal jargon used by lawyers confuses children and places their competence to answer difficult questions on the stand rather than their knowledge of the event in question (Perry, McAuliff, Tam, & Claycomb, 1995). Not only do attorneys in general produce incoherent questions, but also defense attorneys tend to be less supportive and ask more complex and developmentally inappropriate questions than the prosecution (Cashmore & DeHaas, 1992; Davies & Seymour, 1998; Flin, Bull, Boon, & Knox, 1993; Goodman et al., 1992; Perry et al., 1995). Although defense attorneys appear to inappropriately question children, the procedure of cross-examination has been described by the United States Supreme Court as "the greatest legal engine ever invented for the discovery of truth" (e.g., California v. Green, 1970, p. 158). Most recently, the court has emphasized the importance of cross-examination in limiting the admissibility of hearsay when a criminal defendant is not given the opportunity to cross-examine the hearsay declarant (Crawford v. U.S., 2004). Cross-examination allows all parties an opportunity to challenge the evidence of the other party, testing evidence for accuracy and authenticity, helping judges and jury members to determine the truth.

Researchers interested in the interplay between complex questions and child eyewitness' reports have demonstrated that attorneys' language influences the accuracy of a child's testimony (Zajac, Gross, & Hayne, 2003; Zajac & Hayne, 2003). Defense attorneys consistently ask more complex or difficult questions, resulting in poor understanding and lower accuracy in children's responses (Cashmore & DeHaas, 1992; Zajac et al., 2003). Children also rarely ask for clarification and respond simply to ambiguous and nonsensical questions (Zajac et al., 2003). For example, Waterman, Blades, and Spence (2000) interviewed 5- to 8-year-olds using sensible and nonsensible (unanswerable) questions. Results revealed that while children correctly answered sensible questions they incorrectly attempted to answer closed nonsensible questions. One possible explanation for why the defense tends to ask more difficult and complex questions is that they are attempting to undermine the credibility of the child witness (Leippe, Brigham, Cousins, & Romanczyk, 1989). Previous research has indicated that the consistency of children's reports, amount of details disclosed, confidence and projected intelligence influence jury's perceptions of child witness credibility when the evidence presented in the trial is ambiguous (Goodman, Goldings, & Haith, 1984). Since children often have trouble understanding difficult questions (Cashmore & DeHaas, 1992), they may appear less confident and intelligent, and thus less credible, to jury members. However, it is unclear as to whether complex questions, particularly those by the defense, can influence the juror's decisions and lead to a verdict favorable to the defendant.

Forensic linguistics, defined as a sub-discipline of applied linguistics that assesses the interplay between language and the law and crime, is a relatively new area of research, that may be particularly helpful to assess how linguistic complexity affects trial outcome. There is a wide range of research areas in forensic linguistics including the language of legal texts and terminology, the provision of linguistic evidence as well as the language of legal processes (i.e., cross-examinations, interviewing techniques, etc.). Forensic linguistics is also applicable to the assessment of eyewitness veracity (Dulaney, 1982; Friedman & Tucker, 1990; Newman, Pennebaker, Barry, & Richards, 2003).

Coinciding with this increased interest in forensic linguistics was the development of computer-based linguistic analysis software. There has been three main areas of computer-based linguistic software development: (1) automated transcription, (2) word counting and classification software (e.g., Linguistic Inquiry and Word Count Software), and (3) syntax tagging software. Automated transcription transforms voice dictation audio files into written

text. This process of automatic transcription is much faster and less costly than hiring humans to transcribe text for legal and research purposes. Word counting and classification software, such as the Linguistic Inquire and Word Count software, or LIWC (Pennebaker, Frances & Booth, 2001) classifies words into numerous different linguistic dimensions including language categories (e.g., noun, or preposition), relativity related words (e.g., time, or motion), psychological processes (e.g., positive and negative emotions), and traditional contents (e.g., religion or occupation). The total number of words in each category are then counted and percentage of total word scores are used to compare text. LIWC software has been successfully used to compare deceptive and truthful transcribed statements (Newman et al., 2003). Finally, syntax tagger programs, used in the present study, analyze sentences by recognizing part-of-speech classes, such as nouns, adjectives, or verbs and produce noun phrase syntax (see Chaski, 2004; Grant, 2003). Syntax tagger programs have been used to study syntax from many perspectives including the development of language acquisition in children (Parisse & Le Normand, 2000). This advancement of computer-based linguistic analysis opens the door for forensic linguistic research, allowing for fast, consistent and reliable methods to be applied across cases, studies, and laboratories.

The present study utilizes a state-of-the-art linguistic analysis to automatically code and analyse the number of words and complexity of the defense's and prosecution's questions. The purpose of the present study was twofold. First, we examined whether the use of syntax tagger programs would replicate previous findings using human coders. Specifically, whether defense attorneys would use more complex questions when compared to prosecution attorneys. Second, previous studies have neglected to assess whether the complexity of questions asked by the defense and prosecution actually influence the trial outcome. This study addressed this issue by evaluating whether a trial outcome can be predicted based on the complexity of questions asked by either the defense or prosecution.

In the present study, we analyzed 46 transcripts of child sexual abuse trials in California to obtain both complexity and wordiness measures of questions asked by the defense and prosecution. Based on previous findings, we predicted that the defense would use more complex and wordy questions in comparison to the prosecution (e.g., Davies & Seymour, 1998; Zajac et al., 2003). We also expected that more complex language would be related to an acquittal of the defendant as children would appear less credible by jury members when responding to complex questions.

METHODS

Transcripts

Forty-six transcripts from felony child sexual abuse court transcripts held in Los Angeles County, California were obtained. Out of 309 felony child sexual abuse cases that went to trial in Los Angeles between 1997 and 2001, we were able to obtain victim testimony in 243 cases. In 223 of these cases, at least one victim who testified was under 18 at the time. We randomly sampled cases from these 223 transcripts such that (1) half of the cases resulted in a conviction and half in an acquittal and (2) the age of the child witness were matched across verdicts. The mean age of child witness in these transcripts was 11.3 years ($SD = 2.59$, range = 5–15 years) with 11 males (acquitted: 5 males; convicted: 6 males). The mean length of transcripts was $M = 1792.93$ lines, (range of 70–5771 lines).

Coding

Since we were interested in the questions asked by the defense and prosecution directed toward child witnesses, all questions asked by the court or discussions among the

prosecution, defense, and court were eliminated from transcripts. Each transcript was then divided into questions asked by the prosecution and questions asked by the defense. The mean number of questions asked by the defense and prosecution was $M = 239.26$, $SD = 174.43$ and $M = 344.00$, $SD = 216.38$, respectively.

Automated Linguistic Analysis

Connexor Functional Dependancy Grammer (FDG) parser was used to obtain complexity and wordiness measures of the defense and prosecution's questions. The software produced the total number of layers (complexity measure) and branches (wordiness measure) for each question. Then, mean scores for complexity and wordiness for each child witness were calculated for the defense and prosecution.

Connexor FDG parser is a linguistic software program designed to extract linguistic information from natural language text, providing a detailed analysis of syntactic relations between words. Connexor programs have a corpus size of 88 million and recognize 242,000 unique word forms in English. A large corpora from various sources including bureaucratic documents (law text, national/international agreements, etc.) and literature were used to compile the lexicons used in the software analysers. The word class accuracy for the software program in general is 99.3% and the precision of linking subjects and objects is 93.5%.

The Connexor FDG parser builds functionally labeled dependencies between words and assigns morphosyntactic tags to each word (see Järvinen et al., 2004). The software program produces an analysis of each sentence, as denoted by a period. Each sentence is parsed into noun and verb phrases creating a visual tree (see Fig. 1). Each phrase is then further parsed into tokens (i.e., words). Each time a sentence is parsed, it creates a *layer* and each word in the sentence creates a *branch* in the visual tree. The number of layers in each sentence provides a measure of sentence complexity, while number of branches provides a measure of wordiness. For example, the sentence, "Do you recall testifying in April and saying that your mother cleaned up after you threw up?" (Sentence A) produces 4 layers and 10 branches while a more complex sentence such as, "Do you recall telling us that your mother had cleaned up after you throwing up back in April when you testified?" (Sentence B) produces 6 layers and 19 branches (see Fig. 1). The more noun phrases (or layers) used in a sentence, the more complex the sentence is. Similarly, the more words (or branches) in a sentence, the more wordy the sentence is.

RESULTS

The Number of Words and Complexity per Question Asked

Paired sample *t*-tests revealed no significant difference in the number of words used by the defense ($M = 33.88$, $SD = 9.91$) and prosecution ($M = 32.40$, $SD = 19.80$), $t(45) = .88$, $p = .38$. Also, the defense and prosecution question wordiness measures were not significantly correlated with witness' age, $r(46) = .17$, n.s., and $r(46) = .27$, n.s., respectively.

The mean complexity scores of questions asked by the defense ($M = 17.55$, $SD = 3.66$) and prosecution ($M = 16.66$, $SD = 3.66$) were not significantly different, $t(45) = 1.33$, $p = .19$. Again, the defense or prosecutions question complexity scores were not significantly related to age, $r(46) = .18$, n.s., and $r(46) = .27$, n.s., respectively.

Relation Between Verdict and Complexity of Questions

A hierarchical logistic regression was conducted to assess the relation between the complexity of questions and the trial outcome. The trial verdict (convicted vs. acquitted) was

used as the predicted variable, with age entered on the first step, followed by the mean wordiness scores for the defense and prosecution on the second step, and the mean complexity scores for the defense and prosecution on the final step. Wordiness scores were entered second because the number of words used per question and the complexity of questions were significantly correlated $r(46) = .88$, $p < .001$. However, it is possible to ask a complex question using a few words or a simple question using many words. Thus, the wordiness of questions was controlled for in all analyses to clearly assess the relation between complexity and trial outcome.

Neither the first model with age, nor the second model with age and wordiness were significant, $\chi^2(1, N = 23) = .003$, n.s. and $\chi^2(3, N = 23) = 3.92$, n.s. However, the third block with complexity measures was significant, $\chi^2(2, N = 23) = 9.89$, $p < .01$, Nigelkerke $R^2 = .35$, Nigelkerke R^2 change = .24, accurately predicting the trial outcome 73.9% of the time. Further inspection of the final logistic equation revealed that only the mean defense complexity was a significant predictor of verdict ($\beta = .77$, *Wald* = 6.81, $p < .01$, odds ratio = 2.16). Contrary to our prediction, defense lawyers who use more complex questioning were 2.16 times more likely to produce a guilty verdict for their client than those who use less complex questions. In contrast, the prosecution's questions were not significantly related to the trial outcome.

To tease apart the surprising result of defense attorney's complex questions resulting in a conviction, we next examined how children responded to questions asked by the defense. Questions asked by the defense were subdivided into six categories based on children's responses of *I don't know, no, no expansion, yes, yes expansion*, and *open ended* response (see Table 1 for complete descriptions).

A second hierarchical logistic regression was conducted focussing on how the defense's questions affected trial outcome.1 The trial verdict was used as the predicted variable (convicted vs. acquitted). Age entered on the first step, followed by the mean wordiness score for the defense, and finally the mean defense complexity scores for each of the six child response categories (don't know, no, no-expansion, open-ended, yes, and yes-expansion). Neither the first model with age nor the second model with age and mean defense wordiness was significant, $\chi^2(1, N = 23) = .003$, *n.s.* and $\chi^2(2, N = 23) = 2.74$, *n.s.* However, the third block with mean defense complexity was significant $\chi^2(6, N = 23) = 25.94$, $p < .01$, Nagelkerke $R^2 = .62$, Nigelkerke R^2 change = .54. The final prediction equation for the full model was:

$$
\begin{aligned}
\text{logit (outcome)} = \quad &-.175\,(\text{age}) - .60\,(\text{defense wordiness}) \\
&+1.44\,(\text{defense don't know complexity}) \\
&+2.53\,(\text{defense no expansion complexity}) \\
&-2.81\,(\text{defense no complexity}) \\
&+2.42\,(\text{defense open}-\text{ended complexity}) \\
&+.70\,(\text{defense yes expansion complexity}) \\
&+1.62\,(\text{defense yes complexity}) \\
&+2.27, \text{where for outcome} \\
&1 \quad \text{equals conviction and 2 equals acquittal.}
\end{aligned}
$$

The model accurately predicted the trial outcome 82.6% of the time. Specifically, when the defense asked more complex questions leading to an I don't know ($\beta = 1.44$, *Wald* = 6.29, $p < .01$, odds ratio = 4.22) or no-expansion ($\beta = 2.53$, *Wald* = .90, $p < .01$, odds ratio = 12.55) response from children they were more than 4 and 12 times more likely, respectively to receive a conviction verdict than those defense attorneys who used less complex questions leading to such responses.

DISCUSSION

The present study examined the effect of the defense and prosecution attorney question complexity on real world child sexual abuse trial outcomes. Although there was no significant difference in the wordiness or complexity of questions asked by the defense and prosecution, the *complexity* of questions asked by the defense was significantly related to the trial outcome of child sexual abuse cases. Contrary to our prediction, the more complex the defense's questions are, the more likely the trial will result in a conviction of the defendant. In fact, when a defense attorney uses more complex questions, they are over two times more likely to achieve a conviction of their own client in comparison to when they use less complex questions. Conversely, the complexity of questions asked by the prosecution was not significantly related to the trial outcome.

We also found that the complexity of questions asked by defense attorneys was related to the verdict depending on the response given by the child. When complex questions lead to an 'I don't know' or 'no-expansion' response, a conviction verdict was significantly more likely to occur. This suggests that juries may respond positively when children react to defense complex questions in certain ways. Simply responding with 'yes' or 'no' to complex questions was not related to a conviction, but replying with a 'no' and expanding on the response helped achieve a conviction. This may reflect children's ability to successfully resist defense attorney's complex and leading questions. In addition, jury members may perceive a child's response of "I don't know" to a complex question as a sign of competence in their ability to identify questions they do not understand, rather than a lack of memory or a submissive response.

To the best of our knowledge, the present finding is the first to refute the assumption that the complexity of questions asked by the defense undermine the credibility of child witnesses (Zajac & Hayne, 2003; Zajac et al., 2003). There are some possible explanations as to why our results conflict with this assumption. First, jury members may feel the defense's use of complex questions is unfair or a deliberate attempt to mislead the child witness. Perceptions of unjust questioning may lead jury members to feel protective of, or empathetic towards, the child witness. Another possible explanation is that defense attorneys may use more complex questions when the prosecution has a strong case in hopes to mislead or "trick" the child witness, thus more complex questions may occur in cases that result in a conviction verdict. Future research may specifically test these hypotheses. For example, a jury study could manipulate the complexity of questions asked and children's responses to such questions, and assess jurors' verdicts.

There are a few limitations to the present study. First, as there was a larger sample of female child witnesses than males (only 24% males) in the present study, future studies are also needed to assess whether there are gender differences in how question complexity influences jury member's decisions. Second, due to the naturalistic nature of the data random assignment was not possible in this study. Thus, we were not able to control for other variables that may be co-varying, decreasing internal validity. Future experimental studies are needed to address whether other variables may also be driving the results.

The present findings highlight the value of automated forensic psycho-linguistic analyses in assessing the interplay of language between children and adults. Prior to the development of automated linguistic programs, a well-trained linguist was required to code such syntactic information from text. However, the development of automated linguistic software programs allows for this complex linguistic information to be quickly analyzed by non-linguists, allowing this tool to be applied in contexts such as the justice system. This automated linguistic analysis can also be utilized as a professional development tool for attorneys. As

both the defense and prosecution were equally likely to use complex questions and neither adjusted the complexity of their questions according to the age of the child witness, the automated linguistic analysis could be utilized to provide feedback to attorneys on the complexity of their questions and how to revise their questions to make them more developmentally appropriate. Moreover, the automated linguistic analysis might also be useful in training frontline workers who interview children, such as police officers and social workers.

Not only can automated linguistic analyses be applied to the attorney's statements, but also to the witnesses' statements. For example, the LIWC software has been used to identify linguistic markers of deception such that liars tend to have lower cognitive complexity, fewer self-references and other-references and more negative emotion words in their statements compared to truth-tellers (Newman et al., 2003). Additional studies need to be conducted using the Connexor FDG software comparing the complexity of deceptive statements made by witnesses compared to truthful witnesses to gain a greater understanding of whether linguistic complexity may be a marker of deception. In addition, by analyzing the complexity of child witnesses' statements, using automated linguistic software programs such as Connexor FDG, we may be able to gain an understanding of child witness's linguistic maturity. Studies are needed to assess whether the complexity of language used by children themselves is related to the level of complexity of questions at which they can understand. If this is the case automated linguistic software programs can be used to assess children's linguistic maturity and to ensure questions asked of the child match their maturity. Future research assessing the linguistic markers of veracity and developmental markers of linguistic maturity will help shed light on the potential power of forensic linguistics in the practice of law.

REFERENCES

Brennan, M.; Brennan, RE. Strange language: Child victims under cross examination. Riverina Murray Institute of Higher Education; Wagga Wagga, Australia: 1988.

Bruck M, Ceci JS, Hembrooke H. Reliability and credibility of young children's reports. The American Psychologist 1998;53(2):135–151. doi:10.1037/0003-066X.53.2.136.

California v. Green. 399 U.S. 149. 1970.

Cashmore, J.; DeHaas, N. The use of closed-circuit television for child witnesses in the act. 1992. Australian Law Reform Commission Research Paper No. 1

Chaski CE. Forensic linguistics: An introduction to language, crime, and law. The International Journal of Speech, Language and Law 2004;11:298–303.

Crawford v. U.S.. 543 U.S. 817. 2004.

Davies E, Seymour WF. Questioning child complainants of sexual abuse: Analysis of criminal court transcripts in New Zealand. Psychiatry, Psychology and Law 1998;5(1):47–61.

Dulaney EF. Changes in language behavior as a function of veracity. Human Communication Research 1982;9:75–82. doi:10.1111/j.1468-2958.1982.tb00684.x.

Flin RH, Bull R, Boon J, Knox A. Child witnesses in Scottish criminal trials. International Review of Victimology 1993;2:309–329.

Flin RH, Stevenson Y, Davies GM. Children's knowledge of court proceedings. The British Journal of Psychology 1989;80:285–297.

Friedman, HS.; Tucker, JS. Language and deception.. In: Giles, H.; Robinson, WP., editors. Handbook of language and social psychology. John Wiley; New York: 1990.

Goodman GS, Goldings JM, Haith MM. Jurors' reactions to child witnesses. The Journal of Social Issues 1984;40:139–156.

Goodman GS, Quas JA, Bulkley J, Shapiro C. Innovations for child witnesses: A national survey. Psychology, Public Policy, and Law. Special Issue: Hearsay Testimony in Trials Involving Child Witnesses 1999;5:255–281.

Goodman GS, Taub EP, Jones DP, England P. Testifying in criminal court: Emotional effects on child sexual assault. Monographs of the Society for Research in Child Development 1992;57(229):79–80.

Grant T. Forensic Linguistics: Advances in Forensic Stylistics. The International Journal of Speech Language and Law 2003;10:154–156.

Järvinen, T.; Laari, M.; Lahtinen, T.; Paajanen, S.; Paljakka, P.; Soininen, M., et al. Robust language analysis components for practical applications. 2004 [13 June 2008]. from CiteSeer. ITS Scientific Digital Librarywebsite: http://citeseer.ist.psu.edu/726318.html

Leippe, MR.; Brigham, JC.; Cousins, C.; Romanczyk, A. The opinions and practices of criminal attorneys regarding child eyewitnesses: A survey.. In: Ceci, SJ.; Ross, DF.; Toglia, MP., editors. Perspectives on children's testimony. Springer-Verlag; New York: 1989. p. 100-130.

Newman ML, Pennebaker JW, Barry DS, Richards JM. Lying words: Predicting deception from linguistic styles. Personality and Social Psychology Bulletin 2003;29:665–675. doi: 10.1177/0146167203029005010. [PubMed: 15272998]

Parisse C, Le Normand M. How children build their morphosyntax: The French case. Journal of Child Language 2000;27:267–292. doi:10.1017/S0305000900004116. [PubMed: 10967888]

Pennebaker, JW.; Francis, ME.; Booth, RJ. Linguistic Inquiry and Word Count: LIWC 2001. Erlbaum Publishers; Mahwah NJ: 2001. (www.erlbaum.com)

Perry NW, McAuliff BD, Tam P, Claycomb L. When lawyers question children: Is justice served? Law and Human Behavior 1995;19:609–629. doi:10.1007/BF01499377.

Peters WW, Nunez N. Complex language and comprehension monitoring: Teaching child witnesses to recognized linguistic confusion. The Journal of Applied Psychology 1999;84(5):661–669. doi: 10.1037/0021-9010.84.5.661.

Waterman AH, Blades M, Spence C. Do children answer nonsensical questions? The British Journal of Developmental Psychology 2000;18:211–225. doi:10.1348/026151000165652.

Zajac R, Gross J, Hayne H. Asked and answered: Questioning children in the courtroom. Psychiatry, Psychology and Law 2003;10:199–209.

Zajac R, Hayne H. I don't think that's what really happened: The effect of cross-examination on the accuracy of children's reports. Journal of Experimental Psychology: Applied 2003;9:187–195. doi: 10.1037/1076-898X.9.3.187. [PubMed: 14570512]
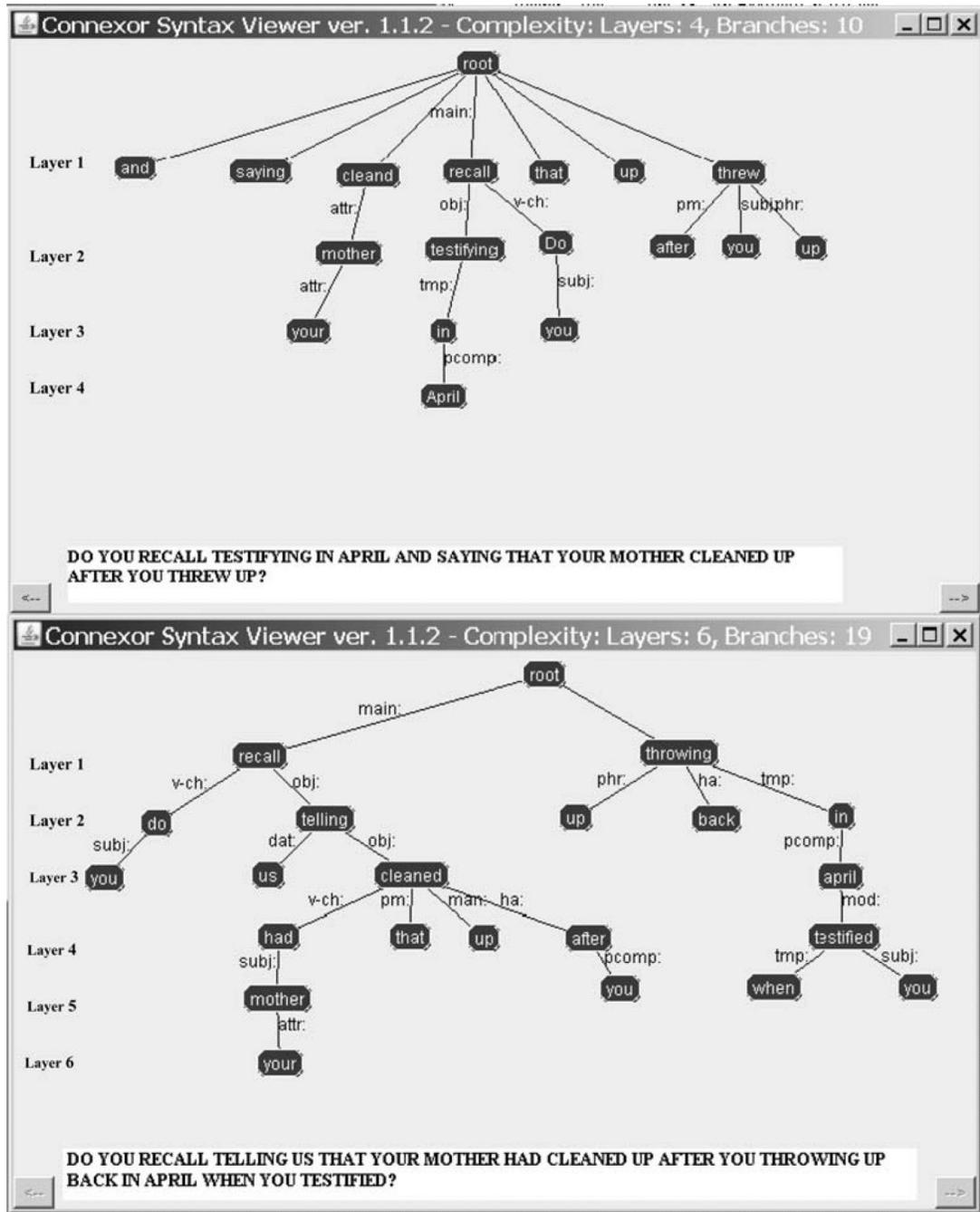
Fig. 1.
The layer and branches count for example sentence A and B

Table 1

Response categories for questions asked by the defense and prosecution

Response type	Defined as	Examples
Don't know	Questions that lead to an "I don't know" response from children.	" I don't know" " I don't remember" " I am not sure"
No	Questions that lead to a "no" response from children.	" No."
No expanded	Questions that lead to a no response plus additional information from children.	" No, but I remember that he grabbed me by the hand."
Yes	Questions that lead to a "yes" response from children	" Yes."
Yes expanded	Questions that lead to a yes response plus additional information from children.	" Yes, that's right. It was a quarter after five and I had just come back from school."
Open ended	Questions that lead to an open-ended response from children.	" It happened 2 months ago when I went to visit my aunt. Mommy said uncle Tom would be out of town. I didn't want to leave my house then."