International Journal of Interpreter Education

Volume 9 | Issue 1 Article 4

June 2017

Using Propositional Analysis to Assess Interpreting Quality

Yan Lydia Ding
University of Auckland

Using Propositional Analysis to Assess Interpreting Quality

Yan Lydia Ding[1]

The University of Auckland

Abstract

This article compares two methods of assessing interpreting quality: the *holistic* method and the proposed *propositional analysis* method. The author first summarizes previous research on interpreting quality, from which quality criteria were selected for holistic assessment. Following that, Turner and Greene's (1978) proposition guideline is briefly introduced as a basis for propositional analysis. Third-year interpreting students were assigned an in-class interpreting task, and their interpreting outputs were recorded, transcribed, and assessed using both methods. Results showed that the two assessment methods agreed with each other in general; however, the propositional analysis method had a few advantages over the holistic assessment method. Propositional analysis gives educators and researchers a clearer overview of the difficulties student interpreters encounter during the interpreting process, by identifying the elements of the source text that were the most challenging for the students. Propositional analysis also facilitates metalinguistic analysis, such as the analysis of different types of propositions and specific language features, so that interpreter educators and researchers can be better informed about the cognitive process involved in interpreting process.

Keywords: interpreting quality, propositional analysis, holistic assessment, quality criteria

[1] *Correspondence to: lydiading84@gmail.com*

Using Propositional Analysis to Assess Interpreting Quality

Assessing interpreting quality has long been a challenge for both interpreting educators and researchers. It is a time-consuming task, and there is no systematic and unified assessment method. As Reiss (2014, p. vi) comments, "The standards… are generally arbitrary". The situation has changed much since Newmark (1982, p. 46) made the assertion on credible translation quality assessment that "detailed schemes for assessing translation are … dead ducks—either too theoretical or too arbitrary". Assessors arbitrarily choose a set of criteria and mark the recorded interpretations (either in the form of transcripts or audio recordings, mostly the latter) against the preselected criteria. This widely adopted method introduces assessors' biases and intuitive judgements. On the one hand, the preselected criteria are, to a large extent, arbitrary and limited in scope; on the other hand, the assessors' holistic judgements are, without doubt, subjective in nature. Such *holistic assessment* generally results in an overall score which represents trainees' global performance, yet educators benefit little from this overall score. They gain little insight into, for example, which part of the source speech poses the most difficulty to trainees and why; so the assessment does not provide information that might help improve trainee competence. Compared with the traditional holistic method, objective *propositional analysis* may be more helpful to educators and researchers. Propositional analysis can detect specific language features, providing educators with valuable information for curriculum focus. It has to be noted, however, the propositional analysis performed in this study assessed the semantic content of students' interpretations only, not the linguistic aspect or delivery.

1. Interpreting Quality

Interpreting quality is the central topic in interpreting studies. After discussing the topic for more than 40 years, researchers still do not agree on the key elements in assessing interpreting quality and on how to accurately measure it (Anderson, 1979; Barik, 1971; Grbić, 2008; Hansen, 2009; Macdonald, 2013; Moser-Mercer, 2008; Pöchhacker & Zwischenberger, 2010). The concept is "elusive" (e.g., Krämer, 2006; Shlesinger et al., 1997), and to some extent, subjective, with the judgement of "excellence" relying much on the assessors' subjective opinions. Nevertheless, researchers have agreed on a few core "linguistic aspects" (Kopczynski, 1994, p. 190), such as "equivalence", "fidelity", and "accuracy" (Pöchhacker, 2002, p. 96), when assessing interpreting quality. Others also propose pragmatic or contextual issues that need to be taken into consideration (Moser-Mercer, 1996).

1.1. Holistic assessment

Subjective assessment of interpreting quality can be reduced if the assessors are experts in the field who rely on their knowledge on a wide range of related domains, including morphosyntactical and microtextual analysis and environmental factors affecting the process. Subjectivity can be further reduced if the assessors apply a consistent set of standards and work in teams of two or more (Williams, 1989). It is natural that different user groups would

have different expectations; that is, scholars and researchers (Mackintosh, 1983; Messina, 2002; Moser-Mercer, 1996; Pöchhacker, 2002; Pöchhacker & Zwischenberger, 2010; Riccardi, 2002; Zwischenberger, 2010) and interpreters and users (Cai & Fang, 2003; Cai & Zeng, 2004; Garzone, 2002; Garzone & Viezzi, 2002; Kopczynski, 1994; Kurz, 1989, 1993, 1994, 2001; Kurz, Basel, Chiba, Patels, & Wolfframm, 1996; Kurz, Pöchhacker, & Zwischenberger, 2008; Marrone, 1993; Pöchhacker, 2001; Rennert, 2010; Vuorikoski, 1993) have different criteria, and each criterion carries different weight.

Drawing on these studies, the following criteria and weight were selected as criteria for holistic assessment of this study. Delivery (accent, pleasant voice, etc.) is intentionally left out in the set of criteria, apart from fluency, which is embedded in linguistic performance. There are two reasons for this decision. First, according to Bühler's 1986 survey, for instance, although delivery is considered in users' or assessors' assessment, the weight assigned to delivery is generally low. Second, in this study, students interpreted from their B language into their A language, therefore, differences in delivery would be minimal. One might include delivery if the direction were from A language to B language, to reflect students' B language competence, a fundamental competence in interpreting. However, as interpreting courses are not linguistic courses, interpreting educators might expect students to have acquired the B language to a satisfactory, if not professional, level when they were admitted into the course.

Table 1. Criteria for holistic assessment.

Semantic Content (80%)	**Linguistic performance (20%)**
Sense consistency, accuracy (50%)	Grammatical correctness (25%)
Terminological adequacy (20%)	Adherence to target-language norms (25%)
Logic, coherence (10%)	Fluency (25%)
Clarity (10%)	Stylistic adequacy (25%)
Completeness (10%)	

1.2. Propositional analysis

Propositional analysis is a detailed, micro-assessment of discourse (Kintsch, 1972; Turner & Greene, 1978) that has a special focus on accuracy. When researchers focus on the accuracy of content, that is, when they conduct error counts (Anderson, 1979; Falbo, 2002; Gerver, 1971; Pym, 1992; Turner, Lai, & Huang, 2010; Vilar, Xu, Fernando D'Haro, & Ney, 2006), they face the issue of determining the meaning unit; this is where proposition comes into play. A *proposition* is the smallest unit that can express a complete meaning, which can be in the form of a word, a phrase, a clause or a sentence. There are three types of propositions: predicates, modifications and connectives (Turner & Greene, 1978). When conducted properly, propositional analysis can provide valuable information.

The holistic assessment approach has the advantage of including as many aspects of the interpreting as the researchers would like to embrace. A holistic score may reflect the comprehensive performance of an interpreter; however, it does not tell much about which part of the source speech causes interpreters the most trouble. The disadvantage of propositional analysis is that it ignores other aspects of the interpretation, such as delivery and presentation, yet it allows the researchers to study the local issues that interpreters might have during the interpreting process. In this study, I compared the two assessment methods and checked the congruity of the two methods in assessing interpreting quality..

2. Method

I originally set out to test the effect of subject knowledge on student interpreters' performance. Interpreting students who had been provided different levels of background knowledge took part in the consecutive interpreting experiment, and their interpreting performance was recorded and analysed. The results of the study involve comparison of the two groups' interpreting quality and their actions taken in the interpreting process. During the analysis process, I found that propositional analysis not only assessed students' performance, by pinpointing the most frequent errors, it could support interpreting educators' teaching. Setting the effectiveness of prior knowledge aside, I instead examined the results of interpreting assessment using the two methods. However, because the raw data and analysis are taken from the experiment as originally designed, there are comparisons between the original two groups (terminology group and portfolio group) studied.

2.1. Participants

Participants were recruited from Beijing University of Foreign Studies. A questionnaire and the pretest of their subject knowledge preselected participants, so that their English competence, interpreting training, interpreting experience, and level of prior knowledge were relatively similar. The final selected participants were 22 native Chinese speakers (two male, 20 female), all undergraduate translation and interpreting majors in the third year of a 4-year BA program.

2.2. Procedure

The experiment followed the research design shown in Figure 1. Participants were randomly assigned to either of two groups, who received different levels of background information before the interpreting task. The terminology group (control group) received a list of terms related to the source speech topic, while the portfolio group (experimental group) received the same list of terms plus a portfolio of background articles.

Propositional analysis

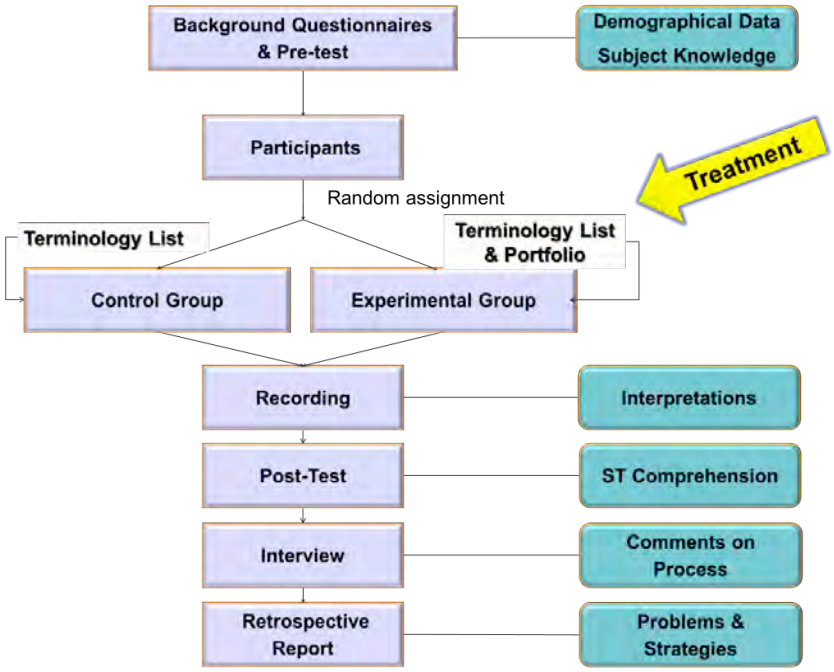

Fig.1. Experiment Design

Participants' interpretations were recorded using the laboratory recording system. Immediately after the interpreting task, all the participants took a post-test, the same as the pretest, to assess whether they had gained more knowledge after interpreting the source speech. Then structured interviews were conducted, with participants invited to comment on the interpreting process. Finally, participants were asked to complete written reports, in which they reflected on their problems and strategies.

2.3. Material

The article chosen for the experiment was published in *The Economist*, titled "Catching a Few More Rays" (2012). It introduces a new type of solar panel and its working mechanism and describes the material used to make it.

Admittedly, written texts have features that are different from speech, such as complex grammar, long sentences, and special vocabulary. However, adopting written texts as source material for interpreting experiments is a common practice in interpreting studies (e.g., Liu & Chiu, 2009; Liu, Schallert, & Carroll, 2004). The source material was adjusted for this study to become more speechlike. Some sentence structures were adjusted, some words were replaced by more colloquial ones and extra connecters were added to be more natural and closer to spoken language. This revised text was manipulated in such a way that the text kept the original logical and structural features of a scientific technical article.

2.4. Assessment

For holistic assessment, two interpreting instructors assessed the participants' interpreting recordings according to the criteria listed in Table 1. Students' recordings were also transcribed and divided into *propositions,* which, according to Turner and Greene (1978), consist of "two or more word concepts … forming a single idea" (p. 2). One hundred propositions were identified in the text. An independent assessor then examined the transcribed interpretations and compared the propositions in the interpretation with the those in the source text. If a proposition in the original text was correctly rendered in the transcribed interpretation, the participant was awarded one point. If the proposition was not adequately reproduced in the target speech, no point was awarded. This analysis allowed the researcher to trace the difficult segments of the source speeches.

3. Results

3.1. Holistic assessment

The two assessors had very high interrater reliability, with a Cronbach's alpha of .898. Table 2 shows that, in general, Assessor 1 tended to give participants higher scores than Assessor 2. As expected, both assessors agreed that participants in the portfolio group performed better than participants in the terminology group. This difference is significant at $p = .01$ (independent-samples t test), and has a large effect size, tested by Cohen's d value.

Table 2. Mean (holistic assessment) of the two groups.

Holistic assessment	**Terminology (control) group**	**Portfolio (experimental) group**	***p*-value**	**Effect size** (Cohen's *d* and effect-size correlation *r*)
Assessor 1	6.087	7.460	.000 **	d =1.79, r = .67
Assessor 2	4.544	6.655	.001 **	d = 1.66, r = .64
Mean	5.315	7.058	.000 **	d = 1.77, r = .66

** $p < .01$.

Participants in the experimental group obtained higher scores for all nine criteria, performing better than the control group in accuracy, coherence, clarity, completeness, fluency, and stylistic adequacy. Their accuracy in terminology adequacy, grammatical correctness, and target language norms also outperformed the control group, but not by as much. This indicates that reading the portfolio of bilingual background articles may have helped participants in the experimental group to better reproduce the source speech with more accurate, coherent, clear, complete, and fluent target speeches.

Propositional analysis

Table 3. Scores for each assessed item.

Holistic assessment	Terminology	Portfolio	Difference	*p*-value	Effect size (Cohen's *d* and effect-size correlation *r*)
Semantic content					
Accuracy	4.5000	6.9091	2.4091	.000 **	*d* = 2.03, *r* = .71
Terminological adequacy	5.7045	6.8636	1.1591	.010 *	*d* = 1.21, *r* = .52
Coherence	5.6136	6.9318	1.3182	.007 **	*d* = 1.28, *r* = .54
Clarity	5.1591	6.8636	1.7045	.001 **	*d* = 1.75, *r* = .66
Completeness	5.8636	7.3182	1.4546	.003 **	*d* = 1.42, *r* = .58
Linguistic performance					
Grammatical correctness	6.9091	7.9773	1.0682	.012 *	*d* = 1.17, *r* = .50
Target-language norms	6.5455	7.4318	0.8863	.023 *	*d* = 1.05, *r* = .46
Fluency	5.2955	6.8636	1.5681	.001 **	*d* = 1.62, *r* = .63
Stylistic adequacy	6.6818	7.8636	1.1818	.003 **	*d* = 1.47, *r* = .59

* $p < .05$. ** $p < .01$.

While holistic assessment is a quick and easy way to assess the relative performance of student interpreters and compare the difference between two groups with different treatments (such as in the original experiment), it does not reveal detailed information such as, for example, which part of the source text poses difficulty to students and is hard to be reproduced accurately, or what language features would challenge student interpreters to produce a satisfactory performance. The result of holistic assessment may also be misleading. If students achieve high marks in terminology adequacy (as did the students in the control group), assessors may be misled into believing that students had mastered the terms quite well, and therefore focus their teaching effort on other aspects, for instance, coherence; yet, incoherence might be directly linked to students' inadequate understanding of terms.

Numerous examples from the control group showed that participants may not have fully understood the terms but nevertheless reproduced the correct equivalence of the terms in the target language.

> ST1: For example, researchers have known for several years that infra-red light can have the same effect on carbon nanotubes.
>
> TT1: 比如说，一些专家发现红外光可以通过碳纳米管传输。
>
> "For example, some researchers found that infra-red light can be transferred through carbon nanotubes."
>
> TT2: 比如说, 红外线已经被利用了碳纳米管来制作电池。
>
> "For example, infra-red light has already been used carbon nanotubes to make batteries."
>
> TT3: 如果将红外光照射在碳纳米管上的话，也能产生电流。
>
> "If we beam infra-red light on carbon nanotubes, electric current can also be generated."

These outputs shows that the terms were reproduced correctly; however, the meaning of the sentence is totally lost; some of the outputs do not make any sense at all. Such incorrect representations prevent educators from learning students' real obstacles in their studies and focusing their teaching accordingly.

Interpreters in the experimental group, by contrast, generated more meaningful and easy-to-understand target texts. They sometimes also provided explanations, or rephrased their own interpretations. For example:

> **TT4:**
> 比如说，几年前我们就已经发现，如果将红外光照射在碳纳米管上的话，也能产生电流。
>
> "For example, we have found several years ago, that if infra-red light shines on carbon nanotubes, it can also generate electrical current."
>
> **TT5:**
> 而经过科学家的研究，其他的材料也可以做到同样的效果。例如说，太阳能光谱当中的红外光就可以通过碳纳米管来发电。
>
> "For example, infra-red light in the sunlight's spectrum can be used to generate electricity through carbon nanotubes."

Propositional analysis highlights the differences in interpreters' terminological adequacy, as well as the fluency and accuracy of their target texts and can be a tool to supplement holistic assessment in identifying students' problems in understanding the source text and delivering the output. In this study, scores from holistic assessment and propositional analysis are in line with each other, cross-validating the two methods (see Table 4).

Table 4. Correlations between scoring of the two assessment methods.

		Holistic assessment	**Propositional analysis**
Holistic assessment	Pearson correlation	1	.823**
	Sig. (two-tailed)	.000	.000
Propositional analysis	Pearson correlation	.823**	1
	Sig. (two-tailed)	.000	.000

Note. Sig. = significance. ** $p < .01$ (two-tailed).

3.2. Propositional analysis

The source speech was divided into 100 propositions, 53 predicates, 31 connectives, and 16 modifications. Table 5 shows that participants in the experimental group achieved significantly higher scores than participants in the control group for all three types of propositions. In addition, participants in the experimental group achieved slightly higher scores for predicates and lower for connectives, whereas participants in the control group obtained higher scores for modifications and lower for connectives.

Propositional analysis

Table 5. Proposition type and propositional scores.

Proposition type	Mean		Standard deviation		*p*-value	Effect size (Cohen's *d* and effect-size correlation *r*)
	Control	Experimental	Control	Experimental		
Predicate	6.019	8.403	3.184	2.107	.000**	$d = 0.883$, $r = .404$
Modification	6.765	8.323	3.276	2.495	.000**	$d = 0.535$, $r = .258$
Connective	5.714	8.071	3.361	2.731	.001**	$d = 0.770$, $r = .359$

** $p < .01$.

The common low score for connectives may indicate that, on a superficial level, these types of propositions were the most difficult for both groups to reproduce. Indeed, to successfully reproduce a connective proposition, one has to have a very good understanding of the preceding as well as the following propositions, so that one can grasp the logic between the sentences before reproducing it in the target language. This can be supported by participants' propositional scores for simple and complex propositions (similar concepts with simple and complex sentences), shown in Table 6.

For both the control and the experimental groups, participants achieved higher scores for the simple propositions than for the complex propositions. Yet this difference is significant for only the control group. This might mean that after reading the portfolio of background articles, participants in the experimental group had a better understanding of the subject matter, so that they could successfully interpret more complex sentences.

Table 6. Proposition type and interpreting quality.

Group	Simple	Complex	*p*-value	Effect size (Cohen's d and effect-size correlation r)
Control	6.809	5.243	.018 *	$d = 0.5091$, $r = .2466$
Experimental	8.556	7.946	.205	$d = 0.2684$, $r = .1330$
p-value	.001**	.000**		

* $p < .05$; ** $p < .01$.

Predicates

Appendix 2 lists the 53 predicates and the number of participants in both groups who successfully reproduced each proposition. The first column of the table shows the number of the proposition, which is also the order in which the proposition appeared in the source speech. The second column lists the actual propositions. The third column, labelled *embedment,* indicates whether the proposition contains embedded propositions as its arguments. The value 1 indicates that, yes, it does contain other propositions as its arguments (the number in the bracket in the proposition represents which proposition is embedded). The value 0 means it does not contain embedded propositions. The fourth and fifth columns show how many participants in each group successfully reproduced the corresponding proposition. The propositions listed in this table are arranged in order from easiest to most difficult, based primarily on the performance of participants in the control group.

Propositional analysis

Appendix 2 also presents the level of difficulty of each individual predicate. Predicates at the top of the table were the most difficult ones, with only a few participants able to correctly reproduce these, whereas the predicates at the bottom of the table were the easiest ones, and almost all the participants in both groups were able to reproduce them correctly. Most of the difficult predicates contain embedment, that is, one or several arguments of these predicates are propositions themselves. Predicates that use other propositions as their arguments increase the difficulty for participants to process information, because participants first need to comprehend the embedded propositions before they can comprehend the main ones (Kintsch & Keenan, 1973; McKoon & Ratcliff, 1980, 2008; Rindflesch & Fiszman, 2003). In addition, the embedded proposition may not be adjacent to the main predicate but a few sentences away, in which case participants would have to recall the earlier information, which increases their mental effort in memorization. Furthermore, the embedded proposition does not always appear in the main predicate as a complete proposition; it may be only a pronoun or another word that functions as a substitute of the embedded proposition. In these cases, participants would have to listen to the speech, comprehend the main and the embedded proposition, recall earlier information, and then create a logical link that connects the embedded proposition and the main predicate. Among these activities, creating logical links may be the most difficult task for participants in the control group, who did not have enough background knowledge on the topic of the source speech.

Modifications

Appendix 3 lists the 16 modifications contained in the source speech. The maximum total score achievable for participants in each group is 176 (16 x 11). However, participants only obtained scores of 103 (control group) and 136 (experimental group). This means participants in the control group reproduced only about 58% of the modifications, and participants in the experimental group reproduced about 77%. In other words, on average, a participant in the control group was able to correctly reproduce nine modifications out of the total 16, whereas a participant in the experimental group was able to reproduce 12. These reproduction rates were slightly higher than the reproduction rates for predicates, for both groups.

While some modifications were difficult to reproduce (for example, Propositions 74, 94, 47, 64, 66, and 58, which fewer than five participants in the control group and fewer than six participants in the experimental group were able to reproduce), other propositions were relatively easy for participants in both groups. The groups contrasted in their reproduction rates for Propositions 40 and 41: Only five participants in the control group correctly reproduced the two propositions, yet all 11 participants managed to reproduce the message accurately. In fact, these two propositions convey the key message in Paragraph 5:

> That discovery led to much experimentation, but little progress. Actually, the chief difficulty lies in the process used to make the **tubes. This process creates** a mixture of two different sorts of tubes: ones that have metal-like properties and ones that are semiconducting. Solar cells need the semiconducting variety. Metallic ones poison the process and must be removed before a cell can work properly.

These two propositions lay the foundation for comprehending the following paragraphs, especially Paragraph 7:

> Dr Strano, however, has exploited a new manufacturing process based on a polymer gel that has an affinity for semiconducting nanotubes, but not for metallic ones. He is thus able to extract large numbers of semiconducting tubes from a mixture…

One of the background articles in the portfolio introduced the two types of nanotubes and how they interact with polymer gels. Perhaps this explains why all participants in the experimental group managed to reproduce the information without effort, whereas participants in the control group were unable to grasp the key message and many failed to reproduce it in the target texts.

The difficult propositions have some common features: (a) they contain no technical terms, and (b) their sentence structures were relatively simple (apart from the fact that they contain embedded propositions). One tentative conclusion in terms of difficulty levels of propositions, therefore, might be that terminology and sentence structure are two factors that affect the difficulty level of individual propositions.

Propositional analysis

Connectives

Connective propositions connect propositions and provide coherence to the text; therefore, the arguments of connectives are, most of the time, also individual propositions. Appendix 4 shows the 31 connectives contained in the source speech, all of which took other propositions as their arguments. Generally speaking, the connectives can be classified into five categories: (a) those such as *and,*" which connect two propositions and which do not have any actually meaning apart from their grammatical function; (b) those such as *but* and *however,* which indicate a contrastive relation; (c) those that express an explanation relation, such as *for example, actually,* and *…means…*; (d) those that express a temporal relation, such as *before, while,* and *after;* and those that express a causal relation, for example, Proposition 71, which indicates that one proposition is the cause of another. Appendix 4 shows that participants in the control group were able to reproduce 59% of all the connective propositions, or 18 connectives out of the total 31. Participants in the experimental group were able to reproduce 73% of all the connective propositions, or 22 connectives out of the total 31. Compared to the other two types of predicates, the difference between the two groups in reproduction rates is the lowest for connectives.

A closer look at Appendix 4 shows that most of incidences of *and* were incorrectly reproduced, meaning that the information was misinterpreted or was not interpreted at all. Yet most of the connectives that indicate comparative or contrastive relations were located at the bottom of the table, which means that most of the participants in both groups were able to reproduce such connectives. This is probably because Chinese texts prefer the use of contrastive conjunctions to connective conjunctions, usually expressed by the ordering of the clauses instead of words that have corresponding grammatical functions (Wen, 2012). Thus most of the participants in both groups chose to not to translate "and", although they translated contrastive connectives appropriately.

4. Discussion and Conclusion

This study compared two methods to assess interpreting quality, holistic assessment and propositional analysis. The two assessment methods complement each other in reflecting students' interpreting products. In addition, the two methods agree with each other, validating each method. While the widely adopted holistic assessment method provides a less time-consuming solution for interpreting educators to monitor students' overall performance and progress, propositional analysis offers interpreting educators a reliable means to examine specific interpreting problems; the results can then be used to guide interpreting teaching.

Although it takes time and considers only the semantic but not the delivery aspect of interpretation, propositional analysis is nevertheless a helpful tool for interpreting educators. By dividing the source text used in the interpreting tasks into individual propositions and then assessing students' reproduction rate of each proposition, interpreting educators can have a direct and visual impression of which propositions were the most difficult ones for students to reproduce, and they can investigate the reasons for the difficulties. Guiding students to use propositional analysis to conduct peer review or self-assessment might save interpreting educators time in the evaluation process, so they can focus their efforts on the pedagogical aspects, for example, designing particular modules to tackle the specific difficulties suggested in the propositional analysis process.

The results of this study demonstrated that simple propositions were easier to reproduce than complex propositions, especially for participants in the control groups. As complex propositions entail complex sentence structures, to understand and reproduce such propositions requires interpreters to go beyond the sentence level and make connections across sentences. Yet, without enough domain knowledge, it may be very challenging for them to do so. Participants in the experimental group, on the other hand, could draw on their prior domain knowledge and "integrate this information into a more complete mental representation of the events […] with minimal reliance on explicit text-based input" (Best, Rowe, Ozuru, & McNamara, 2005, pp. 67–68). In other words, having prior knowledge can help participants process information in a top-down manner, which is more efficient than processing information bottom-up. According to Hawkins (2004), the human brain is a memory-based predicative system that needs to be trained before it can make any inferences. After the brain is provided with information with which to make associated connections or inferences, retrieving information becomes quick and efficient.

Propositional analysis

Deeper propositional analysis, dividing the propositions by type into predicates, modifications and connectives, showed that in general, participants in the experimental groups had higher reproduction rates for predicates and modifications than for connectives. This result represents a direct effect of subject knowledge on information processing; subject knowledge helped participants to select more important information when they were engaged in comprehension. In scientific discourse, predicates (which express the basic ideas in describing action and states) may be the most important type of proposition in constructing ideas—the source speech contained 53 predicates out of a total 100 propositions. Modifications, which by definition "express various forms of restrictions or limitations of one concept by another" (Turner & Greene, 1978, p. 4), may express the logical relations between concepts. These may be less important to comprehension because their role is to modify the basic concepts (there are 16 modifications in the source speech). Finally, connectives (31 in the source speech), which represent the connections between sentences, are the most visible and direct structural signs in a discourse. Each type of proposition plays a different role in discourse, and each one's importance varies according to type of discourse. I argue that in scientific and technical texts, the role of predicates is the most fundamental for comprehension and effective interpreting.

Essential to effective interpreting is the ability to select the most important information in a source text (Liu, Schallert, & Carroll, 2004). The results of this study reflected that participants in the experimental group were able to recognize more important information in the discourse, that is, predicates, and pay less attention to the structural guidance as expressed by the connectives.

One of the reasons for the low reproduction rates of connectives might be that, English and Chinese linking words do not always have a one-to-one relationship. Where a linking word is needed in English discourse, it may be unnecessary in Chinese. The simplified method in assessing connectives the same way as other types of propositions is a limitation of this study, one that came to light only after the analysis was carried out. Future studies are encouraged consider the linguistic features of the two languages involved and optimize the assessment method in rating connectives. In addition, this study only looked at the direction from English to Chinese (B to A); a repetition of the other direction might generate other interesting results.

Propositional analysis makes it possible to detect some of the features in a proposition that made it difficult for participants to comprehend and reproduce. The first such feature is referents, that is, information that has been mentioned earlier in the speech and referred to later in the text, most of the time in the form of anaphora. As discussed, identifying and comprehending referents may be different from and more difficult than comprehending other simple and direct propositions (Burkhardt, 2008), because when trying to understand a referent, one would have to search "working memory for the referent; if [one] does not find the referent in working memory, then [one] searches LTM (long term memory) for an object known as a part of general knowledge" (Kieras, 1977, p. 263). The availability of relevant knowledge is indispensable to successfully understanding a referent, (Frank, Koppen, Vonk, & Noordman, 2007). Because previous research on referents focused on the lower, or lexical, level of the source text (Franceya & Caina, 2014; Nieuwland & Van Berkum, 2008; Pickering & Garrod, 2013; Van Berkuma, Koornneef, Ottena, & Nieuwlanda, 2007; Zwaan, 2014), rather than information processing at sentence and discourse levels, the results of this study can lead to only a tentative hypothesis . More research on referents is needed before any definitive conclusions can be drawn.

Acknowledgments

I would like to thank all the participants who took time from their busy study to participate in this project. Special thanks go to Ms. Wang Xiaoying in Beijing Foreign Studies University, who assisted me in recruiting participants. Last but not least, I would like to express my gratitude to Dr. Ineke Crezee, who encouraged me to publish this article.

Propositional analysis

References

Anderson, L. (1979). *Simultaneous interpretation: Contextual and translation aspects.* Unpublished master's dissertation. Concordia University, Montreal, Canada.

Barik, H. C. (1971). A description of various types of omissions, additions and errors of translation encountered in simultaneous interpretation. *Meta: Translators' Journal, 16*(4), 199–210.

Best, R. M., Rowe, M., Ozuru, Y., & McNamara, D. S. (2005). Deep-level comprehension of science texts. *Topics in Language Disorders, 25*(1), 65–83.

Bühler, H. (1986). Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters. *Multilingua*, *5*(4), 231-235.

Burkhardt, P. (2008). Dependency precedes independency: Online evidence from discourse processing. In A. Benz & P. Kühnlein (Eds.), *Constraints in discourse* (pp. 141–158). Amsterdam, the Netherlands: John Benjamins.

Cai, X., & Fang, F. (2003). Evaluation interpreting quality and effect. *Foreign Languages and Their Teaching, 3*, 41–46.

Cai, X., & Zeng, J. (2004). A review of interpreting performance evaluation studies. *Chinese Translators Journal, 25*(3), 49–54.

Falbo, C. (2002). Error identification and classification: Instruments for analysis. In G. Garzone, P. Mead, & M. Viezzi (Eds.), *Perspectives on interpreting* (pp. 111–127). Bologna, Italy: CLUEB.

Franceya, G., & Caina, K. (2014). Effect of imagery training on children's comprehension of pronouns. *The Journal of Educational Research, 108*(1), 1–10.

Frank, S. L., Koppen, M., Vonk, W., & Noordman, L. G. M. (2007). Modeling multiple levels of text representation. In F. Schmalhofer & C. A. Perfetti (Eds.), *Higher level language processes in the brain: Inferences and comprehension processes* (pp. 133–157). Mahwah, NJ: Lawrence Erlbaum.

Garzone, G. (2002). Quality and norms in interpretation. In G. Garzone & M. Viezzi (Eds.), *Interpreting in the 21st century* (pp. 107–120). Amsterdam, the Netherlands: John Benjamins.

Garzone, G., & Viezzi, M. (2002). *Interpreting in the 21st century: Challenges and opportunities: Selected Papers from the 1st Forlì Conference on Interpreting Studies, 9–11 November 2000*. Amsterdam, the Netherlands: John Benjamins.

Gerver, D. (1971). *Aspects of simultaneous interpretation and human information processing*. Unpublished doctoral dissertation. Oxford University, Oxford, UK.

Gile, D. (1990). L'evaluation de la qualite de l'interpretation par les delegues, Une etude de cas [Evaluation of interpreting quality by delegates – a case study]. *The Interpreters' Newsletter, 3*, 66–71.

Grbić, N. (2008). Constructing interpreting quality. *Interpreting, 10*(2), 232–257.

Hansen, G. (2009). The spec in your brother's eye—the beam in your own: Quality management in translation and revision. In G. Hansen, A. Chesterman, & H. Gerzymisch-Arbogast (Eds.), *Efforts and models in*

interpreting and translation research: A tribute to Daniel Gile (pp. 255–280). Amsterdam, the Netherlands: Benjamins.

Hawkins, J. (2004). *On Intelligence*. New York, NY: Henry Holt and Company.

Kieras, D. (1977). Problems of reference in text comprehension. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 248–269). Mahwah, NJ: Lawrence Erlbaum.

Kintsch, W. (1972). Notes on structure of semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York, NY: Academic.

Kintsch, W., & Keenan, J. (1973). Reading rate and retention as a function of the number of propositions in the base structure of sentences. *Cognitive Psychology, 5*, 257–274.

Kopczynski, A. (1994). Quality in conference interpreting: Some pragmatic problems. In M. Snell-Hornby, F. Pöchhacker, & K. Kaindl (Eds.), *Translation Studies: An Interdiscipline* (pp. 189–198). Amsterdam, the Netherlands: John Benjamins.

Krämer, A.-K. (2006). *Defining the elusive: Interpreting quality research and its pragmatic relevance* Unpublished master's dissertation. The University of Cologne, Bochum, Germany.

Kurz, I. (1989, October 11-14). *Conference interpreting: User expectations.* Paper presented at the 30th Annual Conference of the American Translators Association: Coming of Age, Medford, NJ.

Kurz, I. (1993). Conference interpretation: Expectations of different user groups. *The Interpreters' Newsletter, 5*, 13–21.

Kurz, I. (1994). What do different user groups expect from a conference interpreter? *The Jerome Quarterly, 9*(2), 3–7.

Kurz, I. (2001). Conference interpreting: Quality in the ears of the user. *Meta: Translators' Journal, 62*(2), 394–409.

Kurz, I., Basel, E., Chiba, D., Patels, W., & Wolfframm, J. (1996). Scribe or actor? A survey paper on personality profiles of translators and interpreters. *The Interpreters' Newsletter, 7*, 3–18.

Kurz, I., Pöchhacker, F., & Zwischenberger, C. (2008). *Quality and role: The professionals' view.* Retrieved from http://aiic.net/page/3044

Liu, M., & Chiu, Y.-H. (2009). Assessing source material difficulty for consecutive interpreting. *Interpreting, 11*(2), 244–266.

Liu, M., Schallert, D. L., & Carroll, P. J. (2004). Working memory and expertise in simultaneous interpreting. *Interpreting, 6*(1), 19–42.

Macdonald, P. (2013). It don't mean a thing: Simultaneous interpretation quality and user satisfaction. *The Interpreters' Newsletter, 18*, 35–59.

Mackintosh, J. (1983). *Relay interpretation: An exploratory study.* Unpublished master's dissertation. University of London.

Marrone, S. (1993). Quality: A shared objective. *The Interpreters' Newsletter, 5*, 35–41.

McKoon, G., & Ratcliff, R. (1980). Priming in item recognition: The organization of propositions in memory for text. *Journal of Verbal Learning and Verbal Behavior, 19*, 369–386.

McKoon, G., & Ratcliff, R. (2008). Meanings, propositions, and verbs. *Psychonomic Bulletin & Review, 15*(3), 592–597.

Messina, A. (2002). Quality research and quality standards in interpreting: The state of the art. In G. Garzone, P. Mead, & M. Viezzi (Eds.), *Perspectives on interpreting* (pp. 103–109). Bologna, Italy: CLUEB.

Monitor. (2012, September 1). *Catching a few more rays.* Economics online. http://www.economist.com/node/21560988

Moser-Mercer, B. (1996). Quality in interpreting: Some methodological issues. *The Interpreters' Newsletter, 7*, 43–55.

Moser-Mercer, B. (2008). Constructing quality. In G. Hansen, A. Chesterman, & H. Gerzymisch-Arbogast (Eds.), *Efforts and models in interpreting and translation research* (pp. 143–156). Amsterdam, the Netherlands: John Benjamins.

Newmark, P. (1982). *Approaches to translation.* Oxford, UK: Pergamon Press.

Nieuwland, M. S., & Van Berkum, J. J. A. (2008). The interplay between semantic and referential aspects of anaphoric noun phrase resolution: Evidence from ERPs. *Brain & Language, 106*, 119–131.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*, 329–392.

Pöchhacker, F. (2001). Quality assessment in conference and community interpreting. *Meta: Translators' Journal, 46*(2), 410–425.

Pöchhacker, F. (2002). Researching interpreting quality: Models and methods. In G. Garzone & M. Viezzi (Eds.), *Interpreting in the 21st century: Challenges and opportunities* (pp. 95–106). Amsterdam, the Netherlands: John Benjamins.

Pöchhacker, F., & Zwischenberger, C. (2010). Survey on quality and role: Conference interpreters' expectations and self-perceptions. Retrieved from http://aiic.net/page/3405

Pym, A. (1992). Translation error analysis and the interface with language teaching. In C. Dollerup & A. Loddegaard (Eds.), *The teaching of translation* (pp. 279–288). Amsterdam, the Netherlands: John Benjamins.

Reiss, K. (2014). *Translation criticism-potentials and limitations: Categories and criteria for translation quality assessment.* New York, NY: Routledge.

Rennert, S. (2010). The impact of fluency on the subjective assessment of interpreting quality. *The Interpreters' Newsletter, 15*, 101–115.

Riccardi, A. (2002). Interpreting research: Descriptive aspects and methodological proposals. In G. Garzone & M. Viezzi (Eds.), *Interpreting in the 21st Century* (pp. 15–27). Amsterdam, the Netherlands: John Benjamins.

Rindflesch, T. C., & Fiszman, M. (2003). The interaction of domain knowledge and linguistic structure in natural language processing: Interpreting hypernymic propositions in biomedical text. *Journal of Biomedical Informatics, 36*, 462–477.

Shlesinger, M., Féal, K. D. l., Kurz, I., Mack, G., Cattaruzza, L., Nilsson, A.-L., . . . Viezzi, M. (1997). Quality in simultaneous interpreting. In Y. Gambier, D. Gile, & C. Taylor (Eds.), *Conference Interpreting: Current Trends in Research* (pp. 123–132). Amsterdam, the Netherlands: John Benjamins.

Turner, A., & Greene, E. (1977). The construction and use of a propositional text base (Technical Report No. 63). Boulder: University of Colorado, Department of Psychology. *Institute for the Study of Intellectual Behavior*.

Turner, B., Lai, M., & Huang, N. (2010). Error deduction and descriptors: A comparison of two methods of translation test assessment. *The International Journal for Translation & Interpreting Research, 2*(1), 11–23.

Van Berkuma, J. J. A., Koornneef, A. W., Ottena, M., & Nieuwlanda, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research, 1146*, 158–171.

Vilar, D., Xu, J., Fernando D'Haro, L., & Ney, H. (2006, May 22-28). *Error analysis of statistical machine translation output.* Paper presented at the Fifth International Conference on Language Resources and Evaluation, Genoa, Italy.

Vuorikoski, A.-R. (1993, August 6-13). *Simultaneous interpretation: User experience and expectation.* Paper presented at the Proceedings of the XIIIth World Congress of FIT: Translation—The Vital Link, London, UK.

Wen, Z. (2012). Working memory and second language learning. *International Journal of Applied Linguistics*, *22*(1), 1-22.

Williams, M. (1989). The assessment of professional translation quality: Creating credibility out of chaos. *TTR: traduction, terminologie, rédaction, 2*(2), 13–33.

Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. *Trends in Cognitive Sciences, 18*(5), 229–234.

Zwischenberger, C. (2010). Quality criteria in simultaneous interpreting: An international vs. a national view. *The Interpreters' Newsletter, 15*, 127–142.

Propositional analysis

Appendix 1

Source Text

Catching a Few More Rays

Good morning, ladies and gentlemen, today, I would like to talk about energy technology; in particular, I will discuss a new type of solar panel that can turn infra-red light into electricity.

Solar panels get better and cheaper with every passing year. In one way, though, they are still quite underdeveloped. They work only with light in the visible part of the spectrum. However, 40% of the sunshine that reaches the Earth is in, or very close to, the infra-red, which belongs to the invisible part of the spectrum.

A solar cell that could harvest infra-red light would be a benefit to the solar-power business, but building one has so far proved difficult. Now, however, a group of researchers led by Michael Strano at the Massachusetts Institute of Technology have worked out how to do it.

The most commonly used material to make solar cells is silicon. When sunlight strikes the silicon atoms in an ordinary solar cell, it knocks electrons loose and allows them to flow as an electrical current. Light of other frequencies can do the same trick with other materials. For example, researchers have known for several years that, infra-red light can have the same effect on carbon nanotubes.

That discovery led to much experimentation, but little progress. Actually, the chief difficulty lies in the process used to make the tubes. This process creates a mixture of two different sorts of tubes: ones that have metal-like properties and ones that are semiconducting. Solar cells need the semiconducting variety. Metallic ones poison the process and must be removed before a cell can work properly.

Until now, researchers wishing to do that have been forced to select the semiconducting nanotubes one by one and then sticking them in place with glue. It is possible to make a solar cell this way, but it is time-consuming and expensive. Worse, the chemical instability of the glue means such cells tend to break down rapidly.

Dr Strano, however, has exploited a new manufacturing process based on a polymer gel that has an affinity for semiconducting nanotubes, but not for metallic ones. He is thus able to extract large numbers of semiconducting tubes from a mixture. That done, he deposits them in a thick layer on top of a piece of glass. Their own weight will cause them to stick to the glass without the need for glue. The whole thing is then topped with a layer of buckminsterfullerene, a form of carbon in which the atoms are organized as spheres. This buckminsterfullerene layer acts as an electrode, and conducts away the electricity produced by the nanotubes

The result is not exactly efficient. The cell transforms only around 0.1% of the infra-red light thrown at it into electricity (compared with 20% for an ordinary solar cell). But Dr Strano and his colleagues are excited about the result. After all, 0.1% is a big step up from nothing at all, and most existing solar technologies began with similarly poor efficiencies that were improved gradually over the course of time.

Moreover, the new technology has one big benefit. Though the carbon nanotubes absorb infra-red light, they are almost totally transparent to the visible variety. This means that, if and when they become commercialized, they can be overlaid on traditional silicon cells. This new device will convert a larger fraction of the incoming sunlight into electricity. Thank you.

Appendix 2

Predicates

No	Proposition	Emb[1]	Ctrl	Exp
81	throw (, infra-red light [80], at the cell)	1	0	3
78	produce (the nanotube, electricity [77])	1	0	4
67	is able to (Dr Strano, [68])	1	0	5
70	deposit ([68], the semiconducting tubes, in a thick layer on top of a piece of glass)	1	1	6
68	extract (Dr Strano, large numbers of semiconducting tubes from a mixture)	0	1	7
39	creates (this process [38], a mixture of two different sorts of tubes)	1	2	5
72	is topped with (the whole thing [70], a layer of buckminsterfullerene)	1	2	7
73	is (buckminsterfullerene, a form of carbon)	0	2	7
77	conducts away (buckminsterfullerene layer, the electricity [78])	1	2	7
26	knocks ([25], electrons, loose)	1	2	8
63	base ([62], a polymer gel)	1	3	5
98	can be overlaid on (new technology, traditional silicon cells)	1	3	7
43	poison (metallic tubes, the process [38])	1	4	4
75	acts as (this buckminsterfullerene layer, an electrode)	0	4	6
88	began with (most exciting solar technologies, similarly poor efficiencies)	0	4	7
45	remove (, metallic tubes)	0	4	8
32	can have (infra-red light, the same effect [26–28], on carbon nanotubes)	1	4	10
29	can do (light of other frequencies, the same trick [26–28], with other materials)	1	5	8
62	has exploited (Dr Strano, a new manufacturing process)	0	5	8
38	make (process, tubes)	0	5	9

[1] Emb: embedment. The value of 0 means the proposition does not contain embedment, whereas the value of 1 means the proposition contains embedment, represented as the number in brackets.

Propositional analysis

No	Proposition	Emb[1]	Ctrl	Exp
37	lies in (the chief difficulty, the process [38])	1	5	10
93	absorb (the carbon nanotubes, infra-red light)	0	5	10
21	lead (Michael Strano, a group of researchers)	0	6	6
100	Thank you.	0	6	7
25	strikes (sunlight, silicon atoms in an ordinary solar cell)	0	6	8
54	make (, a solar cell this way [50–52])	1	6	9
6	turn ([5], infra-red light, electricity)	1	6	10
31	have known (researchers, that [32])	1	6	10
80	transforms (the cell, 0.1% of the infra-red light, into electricity)	0	6	11
13	reach (sunshine [12], Earth)	1	7	8
99	will convert (this new device, a larger fraction of the incoming sunlight, into electricity)	0	7	9
15	would be ([16], benefit to the solar power business)	1	7	10
35	led to (that discovery [32], little progress)	1	7	11
97	become (the new technology, commercialized)	0	8	8
33	led to (that discovery [32], much experimentation)	1	8	9
49	have been forced to (researchers, [50–52])	1	8	9
12	is, or close to (40% sunshine, infra-red)	0	8	10
14	belongs (infra-red light, invisible part of the spectrum)	0	8	10
16	harvest (a solar cell, infra-red light)	0	8	10
28	allows (electrons, flow as an electrical current)	0	8	10
42	need (solar cells, semiconducting tubes)	0	8	10
91	has (the new technology, one big benefit)	0	8	10
22	belongs ([21], Massachusetts Institute of Technology)	1	9	8
50	select (researchers, the semiconducting nanotubes one by one)	0	9	8
89	improve (the poor efficiencies [88], gradually over the course of time)	1	9	11
52	stick (researchers, semiconducting nanotubes, in place with glue)	0	10	9
23	Is (the mostly commonly used material to make solar cells, silicon)	0	10	11
82	transforms (an ordinary solar cell, 20%)	0	10	11
10	work with (the [7], light in the visible part of the spectrum)	1	11	9
1	Good Morning, ladies and gentlemen	0	11	11
3	would like to talk about (I, Energy technology)	0	11	11

Propositional analysis

No	Proposition	Emb[1]	Ctrl	Exp
5	discuss (I, a new type of solar panel)	0	11	11
20	have worked out (a group of researchers, how to do it (building [16])	1	11	11
Total			317	447
Average number of propositions correctly reproduced per participant			28.8	40.6
Percentage of predicates correctly reproduced			54.3%	76.6%

Appendix 3

Modifications

No	Proposition	Emb	Ctrl	Exp
74	are organized as (the atoms of buckminsterfullerene [73], spheres)	1	1	5
94	are (they [93], totally transparent to the visible variety)	1	2	4
47	can work (a cell [16], properly)	1	3	6
64	has (a polymer gel [63], an affinity for semiconducting nanotubes)	1	3	6
66	has not (a polymer gel [63], an affinity for metallic nanotubes)	1	3	6
58	(the chemical instability of the glue [52]),	1	4	5
40	have (tubes, metal-like properties)	0	5	11
41	have (tubes, semiconducting properties)	0	5	11
60	tend to (such cells [54], break down rapidly)	1	8	9
18	has proved (building [16], difficult)	1	8	10
86	is (0.1% [80], a big step up from nothing at all)	1	9	10
7	get (solar panel, better and cheaper, with every passing year)	0	10	10
9	are (they [7], still quite underdeveloped)	1	10	10
56	is (it [54], time-consuming and expensive)	1	10	11
79	is not (the result, efficient, compared with [82])	1	11	11
84	are (Dr Strano and his colleagues, excited about the result [80])	1	11	11
Total			103	136
Average number of propositions correctly reproduced per participant			9.36	12.36
Percentage of modifications correctly reproduced			58.5%	77.2%

Appendix 4

Connectives

No	Proposition	Emb	Ctrl	Exp
44	And	1	2	2
92	though (93, 94)	1	2	3
69	after (68, 70)	1	2	5
27	And	1	2	8
36	Actually	1	3	5
65	But	1	3	5
76	And	1	3	5
61	however,	1	4	4
46	before (47, 45)	1	4	6
30	for example	1	4	7
87	And	1	4	8
71	will cause (their [68] own weight, tubes, to stick to the glass without the need for glue.)	1	4	9
95	means ([92], that [96])	1	4	9
53	Is (it, possible that [54])	1	6	7
59	means that (60)	1	6	9
24	when (25, 2628)	1	7	8
34	But	1	7	11
96	when (97, 98)	1	7	11
51	And	1	8	9
57	worse,	1	9	8
19	however,	1	9	9
90	Moreover	1	9	9
8	in one way though	1	10	10
11	however,	1	10	10
48	until now, wish to (researchers, [45])	1	10	10

Propositional analysis

55	But	1	10	10
17	But	1	10	11
4	in particular,	1	11	10
85	after all	1	11	10
2	Today	1	11	11
83	But	1	11	11
Total			203	250
Average number of propositions correctly reproduced per participant			18.45455	22.72727
Percentage of predicates correctly reproduced			59.5%	73.3%