Reading Comprehension of Different Genres: A Fuzzy Approach

Amir Hossein Shahballa
Shiraz University, Shiraz, Iran
Tel: 98-361-422-3118   E-mail: ashahballa@yahoo.com

Farshad Alamdar Youli
Shiraz University, Shiraz, Iran
E-mail: falamdaryouli@gmail.com

Received: November 20, 2011     Accepted: December 13, 2011     Published: February 1, 2012
doi:10.5539/ijel.v2n1p17          URL: http://dx.doi.org/10.5539/ijel.v2n1p17

Abstract
This study presents a model based on fuzzy logic with two inputs: idea units and main ideas. The inputs were gained by scoring reading comprehension of 19 MA students at Shiraz University who were each given three different genres. The data were fed into the fuzzy system and fuzzy scores were obtained as the output. First, the same participants’ understandings across different genres and then, different participants’ understandings across the same genre were compared. And finally, the number of idea units referred to by participants was calculated to see which idea units were problematic. The results of comparing fuzzy scores and scores resulting from idea units indicated that fuzzy scores were fairer, and more accurate. The results of paired t-test showed that the narrative genre was easier than the descriptive and argumentative genres. Also, the results showed that the same participants had different degrees of understanding across genres and different participants had different degrees of understanding in the same genre.

Keywords: Fuzzy logic, Understanding, Genre, Reading comprehension, Text difficulty

1. Introduction
According to Harris (2000) the problems we deal with are of three categories. The first category is solved using classical (Aristotelian) logic based on sets with sharp boundaries. In another category of the problems, statistical methods are applied to account for the random uncertainty inherent in the problems. Here the concept of probability is used. However, in many real problems, the uncertainty is neither random nor of a statistical nature. Here the corresponding concept is possibility rather than probability. It is for this last category that fuzzy logic is useful.

The name "fuzzy set", according to Iskander (2007), implies two distinctive characteristics. First, a member belongs to a fuzzy set with a degree of certainty/belief (i.e. a value from 0 to 1) rather than crisp membership in which an element is or is not a member of a set. Second, partial membership is allowed; that is, the boundaries between the sets are not absolute and thus the same element can be a member of more than one set with different membership degrees.

As Zadeh (2008), the pioneer of fuzzy logic, says, “There are many misconceptions about fuzzy logic. Fuzzy logic is not fuzzy. Basically, fuzzy logic is a precise logic of imprecision and approximate reasoning.” (p. 2751). More support for fuzzy logic comes from what Cohen, Manion and Morrison (2000) say:

Fuzzy logic enables us to gain a more precise measurement of the variance within and between these semantic categories; it recognizes that imprecision, rather than bivalence (either something is or is not the case) is a characteristic of many phenomena. Fuzzy logic opts for shades of grayness rather than black-or-white (Kosko, 1994:102)! (p. 389)

McNeill and Thro (1994) clarify very nicely the above quotation with a simple example, using a bowl of oranges! Is the following figure a bowl of oranges? The answer is No.
What about the next figure? Is it a bowl of oranges?

The answer is yes. As McNeill and Thro (1994) say, this is an example of crisp logic. The bowl contains *either* totally apples *or* totally oranges. However, as they truly say, life is more complex. Consider for example the following figure in which someone has swapped one orange for one apple in the apples bowl. Is it a bowl of oranges?

Here, as they say:

The situation itself makes *either* Yes *or* No inappropriate. In fact, if you had to say Yes or No, your answer would be less precise than if you answered One, or Some, or A Few, or Mostly—all of which are fuzzy answers, somewhere in between Yes and No. They handle the actual ambiguity in descriptions or presentations of reality. (p. 6)

Imprecision is, in fact, a characteristic of our understanding of the world phenomena, including understanding of texts. When individuals read a piece of text, it cannot be claimed that they have precisely grasped everything the text conveys; nor can we say they have understood nothing. Rather we can say their understanding is poor, not bad, fairly well, well, etc. Moreover, when different individuals read the same text, they differ in the degree of their understanding.

1.1 Fuzzy Logic

Fuzzy set theory was introduced in a seminal work by Askar-Zadeh, in 1965 (Shams, 1379/2000). Today, fuzzy logic is widely used in many industries following the initial uses made of it by the Japanese (Poursaeed, 1379/2000)

According to Menhaj and Nassaji (1379/2001), human reasoning is too complicated to be described by logical symbols. Furthermore, due to the limitations inherent in classical logic, its symbols are incapable of formulating human reasoning. However, fuzzy logic is different. According to Sarmad-Saeedi (1380/2001), fuzzy logic guides the researcher from the true-false Aristotelian logic in the world of human and social relations to another way of looking at the realities in which we cannot draw a definite line between true and false. (p. 37)

With reference to Figure 1, the MATLAB software help section maintains, “Fuzzy logic is all about the relative importance of precision. How important is it to be exactly right when a rough answer will do?”

1.1.1 Applications of Fuzzy Logic

The idea of fuzzy logic has been greatly used in expert systems and as Cox and O'Hagan (1994) state, "fuzzy expert systems model the world in terms of the semantics associated with the underlying variables, thus providing a much closer relationship between real world phenomena and computer models." This is the reason why this logic gains momentum every day.
Though fuzzy logic is widely associated with engineering fields of study and as Cox and O’Hagan (1994) state, “nearly all the attention that fuzzy logic receives has been in the area of process and control engineering” (p. 2), it has also been used in many social areas. Baba, Kuscu and Han (2009) give some examples of the use of fuzzy logic in social sciences (Kwok, Ma, & Vogel, 2001; Rasmani & Shen, 2005; Feng, Rozenblit & Hamilton, 2008; and Biswas, 1995). Except for Feng, Rozenblit and Hamilton (2008) who used fuzzy logic based performance assessment in the virtual assistive surgical trainer, other studies (Kwok, Ma, & Vogel, 2001; Rasmani & Shen, 2005; and Biswas, 1995) are related to classroom assessment. Also, Baba, Kuscu and Han (2009) developed a software that can be used in multi-purpose decision-making processes such as assessing projects, journals, the performance of teachers, students etc.

Since it is related to measurement, fuzzy logic has a wide use in language testing. For example, Cin and Baba (2008) developed a user-friendly multiple purpose fuzzy multi-criteria assessment software to evaluate the English proficiency of students. Another example is the study done by Baba, Bakanay, and Cin (2009) which indicates that fuzzy logic increases the “quality of education facilities, motivation, reliability, consistency, and objectivity of the assessment” (p. 1). Al-Hammadi and Milne (2004) proposed a Neuro-Fuzzy model to predict student performance before admission to the college. They also mentioned some studies using fuzzy logic in education assessment (Nolan, 1998a & 1998b; Ma & Zhou 2000; Echauz & Vachtsevanovs, 1995; Ranjit, 1995). Another study showing the application of fuzzy logic is Zahedi (2001) in which the partial knowledge of the participants in MC tests has been measured through fuzzy logic. And finally, Zadeh (2008) speaks of what is not widely recognized namely, that one of the most important contributions of fuzzy logic is its high power of precisation. It is this power that underlies the ability of fuzzy logic to serve as a co intensive model of reality, especially in human-centric fields such as economics, law, linguistics and psychology. (p. 2276)

1.2 Reading Comprehension Tests

Hickey (1983) introduces two types of questions: those that teach and those that test reading comprehension. The testing questions can give us a classification of the current test formats used for reading comprehension. They are as follows:

1). Open questions/comprehension questions
2). Multiple-choice questions - to test vocabulary - to test structures
3). True/False questions
4). Complete the sentence” to describe an event in the text.
5). Fill in the blanks ....and many more. (pp. 46-7)

In fact, these are the most common test formats used for testing reading comprehension. All these test formats share one disadvantage which is looking at assessment through black-and-white spectacles. In other words, they either test in a binary fashion i.e. the answer is either correct or wrong (like multiple-choice questions), or are scored in a binary fashion (like the open-ended questions) though the nature of the answer may not be binary.

However, as Al-Hammadi and Milne (2004) state, “Fuzzy set theory is an efficient and effective method to represent the uncertainty and fuzzy terms in the assessment environments” (p. 837), and hence, the fuzzy scores have some advantages over the scores gained by traditional test formats Fuzzy scores are not based on correct-or-wrong answers, rather they are based on a degree of the examinee’s knowledge; they are fair and very similar to the way we assess in reality

1.3 Units of Analysis

A review of the literature reveals that different studies have used different units for analysis (Budd, 1988; Distefano and Valencia, 1980; Andoline, 1980; Bardovi-Harlig and Bozman 1988; Bardovi-Harlig, 1992; Hosseinchary and Yosef 1387/2008; Samim-Banipashemi 1371/1992; Frid, Horne, and Roll 2007; Rashidi and Shahballa 2010). The present study has focused on idea units and main ideas. Idea units are very important in this study since they are used as criteria to score the writings of the participants which reflect their understanding of the texts. Chafe (1985) mainly defines idea units for spoken language. Gee (1996), with reference to Chafe’s (1985) definition, mentions that the “vast majority of these idea units are a single clause, with one piece of new information towards the end of the clause. It is only when the subject of the clause, or an adverbial element, is new information that it constitutes an idea unit by itself” (p. 103).

In this study, idea units are largely synonymous with propositions. Therefore, the definition of propositions can stand for the definition of idea units. Richards (1983) defines propositions as “the basic units of meaning
involved in comprehension” (p. 220). Furthermore, “A proposition is that part of the meaning of a clause or sentence that is constant, despite changes in such things as the voice or illocutionary force of the clause.” (n.n., 2004). As an example, the sentence “the tall open-minded student had an accident” has the following three propositions: 1) The student is tall, 2) The student is open-minded, 3) The student had an accident.

2. Statement of the Problem

Assessment is a very important part of the teaching-learning process and many attempts have been made to offer an objective assessment congruent with reality. Thus, multiple-choice (MC) tests have been devised and widely used due to their objectivity being easy to administer and score. However, they have some disadvantages. For example, guessing has been a great problem in the area of MC tests. Students may have no knowledge of the item, but still have a chance of marking the correct response, or they may know a part of the answer but fail to mark the key and get no credit. Furthermore, they may have partial knowledge of the material and by deleting the infeasible answers get full scores. These problems indicate that MC tests do not reflect the way we really evaluate something in our everyday life. The solution may lie in assessing through fuzzy logic. Indeed, human reasoning is done in terms of fuzzy sets, not of discrete symbols and numbers. These fuzzy terms define general categories, but not rigid, fixed collections. The transition from one category--concept, idea, or problem state--to the next is gradual with some states having greater or lesser membership in the one set and then another (Cox & O'Hagan, 1994, p. 3).

3. Research Questions

This study is concerned with the use of fuzzy logic in assessment. In fact, it presents a fuzzy model with two inputs, main ideas (MIs) and idea units (IUs), and one output, a fuzzy score (Fscore). The inputs are obtained from each text and entered into the model to obtain Fscores which indicate the degree of understanding of the participants. Thus, the study seeks answers to the following questions:

1). How can we assess people’s understanding of texts through fuzzy logic?
2). Do different people have different degrees of understanding from the same genre?
3). Do different sexes have different degrees of understanding from the same genre?
4). Do the same people have different degrees of understanding when they read different genres?

4. Design and Method of the Study

4.1 Participants

Nineteen Iranian MA students of TEFL at Shiraz University ranging in age from 23 to 27, volunteered to participate in this study. In order to have similar conditions for all participants, it was attempted to choose participants from the same educational level.

4.2 Materials

Three texts of different genres, descriptive, narrative and argumentative, were selected. The descriptive and argumentative texts were chosen from Paragraphs and Essays (Emami, 2005). The narrative text was selected from Paragraph Development (Arnaudet & Barrett, 1990). In selecting texts, it was attempted to choose texts at or below the English proficiency level of the participants.

To make a fuzzy model of scoring, the Fuzzy Logic Toolbox of the computer software MATLAB (version 7.1, 2005) was used.

4.3 Procedures

The participants received the texts at different times and were asked to perform the task in class. To maintain equal conditions in performing the task, all the participants could use their dictionaries, but they were requested not to consult anybody for help. Also, they had the choice to write their understanding in either Persian or English. The task had been clarified through examples on the question paper. The task was twofold. Once, the students were supposed to write their understanding unit by unit. Then, they had to write the main ideas of the texts. The data were collected in three different sessions, one genre at a time.

As the goal was to measure the extent to which the participants had understood the meaning present in the text, a unit of measurement was needed. For this reason, idea units (IUs) and main ideas (MIs) were considered as the units of measurement.

The idea units were extracted from the main texts as the criterion for the IU-section. Then, the writings of the participants were scored based on the extracted criterion IUs by the first rater. To check the reliability of scoring,
21 writings out of the 57 available writings were randomly chosen. Then, the sample was given to a second rater. After getting the scores, a Pearson correlation was run and $r = .94$ (p<.001) was obtained. The correlation coefficient was put in Spearman prophecy formula ($\frac{2r}{1 + r^2}$) to calculate the inter-rater reliability. The result was .968. This showed that the rating was highly reliable.

The criterion for scoring the MIs was extracted from the idea units referred to by three Shiraz University instructors as main ideas of the texts. Then, for each idea unit a weight of two was considered. Therefore, the units referred to by the three instructors had a weight of six, those referred to by two instructors had a weight of four and those referred to by only one instructor had a weight of two. On the whole, the descriptive genre consisted of 32 IUs and 22 MIs, the narrative text had 31 IUs and 22 MIs, and the argumentative genre contained 12 IUs and 26 MIs. Then, according to the criteria the writings of the participants were codified and scored.

4.4 Fuzzy Inference Process

This section, based on the help section of MATLAB software, shows the procedure of applying fuzzy logic in the present study. The result is a 2-input, 1-output, 9-rule system. The basic structure is as displayed in figure 2.

Figure 2 shows how the information enters the system from the left. Then, the rules are formed, the results are combined and distilled and finally, a crisp output is produced. According to the help section of MATLAB software there are five stages in the fuzzy inference process:

1). fuzzification of the input variables,
2). application of the fuzzy operator (AND or OR) in the antecedent,
3). implication from the antecedent to the consequent,
4). aggregation of the consequents across the rules, and
5). defuzzification.

Having gone through the above stages, a fuzzy system was defined in the Fuzzy Logic Toolbox section of the MATLAB software, which gets two scores as its inputs (MI and IU) and provides us with an output which is an Fscore. Figure 3 shows the inside of the system in which each one of the inputs has three membership functions: poor, average, and good and the output has five membership functions: very low, low, average, high, and very high.

The user can enter the MI and IU scores in the window “rule viewer” and get the Fscore as shown below. In the following window, for example, the MI=30 and IU=80 have been entered and hence the Fscore equals 54.7.

5. Results and Discussion

To assess understanding fairly and more accurately, a fuzzy system was defined. The MI and IU scores of the participants for each genre were calculated in percentage and entered into the Fuzzy-Scorer to get the final crisp score (or Fscore). The results are shown in Table 1.

As can be seen from Table 1 and based on the explanations provided above concerning the fuzzification and defuzzification of scores, one can conclude that it is possible to assess people’s understanding of texts through fuzzy logic. This can provide a positive answer to our first research question.

To answer the second question, “Do different people have different degrees of understanding from the same genre?” we consider the standard deviation of the Fscores obtained for different genres. The results are presented in Table 2.

The most dispersion is seen in the narrative text and the least in the argumentative; however, different people perform differently when they read the same genre. This can be seen when we consider the chart in Fig. 5. It is only in the argumentative text that some participants have scored similarly. This is seen in the horizontal line in front of score 50.

To answer the third research question concerning the effect of sex on the degree of comprehension, some independent samples t-tests were run. The results are reported in Table 3.

As Table 3 indicates, sex had no effect on the participants’ Fscores in any of the three genres.
To see whether the participants’ understandings differed across genres, a paired samples t-test was run. There were three comparisons. Therefore, according to Bonferroni adjustment, the significance level was determined at .017 (.05/3=.017). The results in Table 4 were obtained.

As evident in Table 4, only the difference between the narrative genre and the argumentative genre was significant. The mean differences for the other two genres were not significant. This indicates that participants have performed equally in the narrative and descriptive genres.

6. Conclusion

The theoretical framework used in this study is fuzzy logic. This framework was used in order to present a model which is fairer, more realistic, and more accurate. This model can, in fact, be a meaningful model for scoring the reading comprehension of the students since it does not rely only on the superficial knowledge of the text; rather it measures their understanding of the purpose of the text (its main ideas) as well. This is very important since students may be able to answer the questions based on what they see on the surface (words and grammar); however, through this model a deeper understanding is considered for scoring as well.

Different text types (genres) had different levels of difficulty. This was evident from the results of paired t-tests which showed a higher mean for the narrative genre than the descriptive and argumentative genres. This shows that the narrative genre is the easiest genre.

Acknowledgement

Our deep gratitude is due to Dr Mortaza Yamini of Shiraz University, Shiraz, Iran, for his very insightful ideas and very generous assistance through this study from the very beginning to the very end. We wish him the best.

References


Table 1. Fuzzy score for each participant’s writing

<table>
<thead>
<tr>
<th>Code</th>
<th>descriptive</th>
<th>narrative</th>
<th>argumentative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MI, IU Fscore</td>
<td>MI, IU Fscore</td>
<td>MI, IU Fscore</td>
</tr>
<tr>
<td>1</td>
<td>18.181, 71.875</td>
<td>36.363, 90.322</td>
<td>23.076, 100</td>
</tr>
<tr>
<td>2</td>
<td>27.272, 50</td>
<td>90.909, 67.741</td>
<td>23.076, 83.333</td>
</tr>
<tr>
<td>3</td>
<td>36.363, 78.125</td>
<td>90.909, 93.548</td>
<td>53.846, 100</td>
</tr>
<tr>
<td>4</td>
<td>27.272, 34.375</td>
<td>48.387, 25.6</td>
<td>0, 33.333</td>
</tr>
<tr>
<td>5</td>
<td>0, 40.625</td>
<td>54.545, 80.645</td>
<td>23.076, 75</td>
</tr>
<tr>
<td>6</td>
<td>18.181, 81.25</td>
<td>90.909, 100</td>
<td>92.307, 100</td>
</tr>
<tr>
<td>7</td>
<td>36.363, 81.25</td>
<td>27.272, 29.032</td>
<td>23.076, 75</td>
</tr>
<tr>
<td>8</td>
<td>63.636, 71.875</td>
<td>63.636, 67.741</td>
<td>23.076, 91.666</td>
</tr>
<tr>
<td>9</td>
<td>27.272, 71.875</td>
<td>72.727, 51.612</td>
<td>0, 66.666</td>
</tr>
<tr>
<td>10</td>
<td>9.090, 93.75</td>
<td>81.818, 100</td>
<td>23.076, 100</td>
</tr>
<tr>
<td>11</td>
<td>0, 59.375</td>
<td>54.545, 80.645</td>
<td>23.076, 100</td>
</tr>
<tr>
<td>12</td>
<td>81.818, 87.5</td>
<td>100, 87.996</td>
<td>23.076, 100</td>
</tr>
<tr>
<td>13</td>
<td>27.272, 56.25</td>
<td>27.272, 90.322</td>
<td>91.666, 50</td>
</tr>
<tr>
<td>14</td>
<td>36.363, 56.25</td>
<td>63.636, 80.645</td>
<td>23.076, 91.666</td>
</tr>
<tr>
<td>15</td>
<td>63.636, 53.125</td>
<td>36.363, 83.870</td>
<td>76.923, 58.333</td>
</tr>
<tr>
<td>16</td>
<td>72.727, 75</td>
<td>81.818, 74.193</td>
<td>23.076, 91.666</td>
</tr>
<tr>
<td>17</td>
<td>72.727, 87.5</td>
<td>63.636, 80.645</td>
<td>69.230, 66.666</td>
</tr>
<tr>
<td>18</td>
<td>63.636, 84.375</td>
<td>54.545, 96.774</td>
<td>23.076, 100</td>
</tr>
<tr>
<td>19</td>
<td>54.545, 84.375</td>
<td>36.363, 25.806</td>
<td>23.076, 66.666</td>
</tr>
</tbody>
</table>

Table 2. Mean Fscores and SD

<table>
<thead>
<tr>
<th>Genre</th>
<th>Mean Fscore</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>68.5263</td>
<td>24.55038</td>
</tr>
<tr>
<td>Descriptive</td>
<td>56.3421</td>
<td>23.29133</td>
</tr>
<tr>
<td>Argumentative</td>
<td>53.6211</td>
<td>15.94245</td>
</tr>
</tbody>
</table>

Table 3. Descriptive statistics and t-test results showing sex effect on Fscores

<table>
<thead>
<tr>
<th>sex</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>t</th>
<th>df</th>
<th>Sig (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>descriptive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>8</td>
<td>45.3125</td>
<td>19.19505</td>
<td>-1.880</td>
<td>17</td>
<td>.077</td>
</tr>
<tr>
<td>female</td>
<td>11</td>
<td>64.3636</td>
<td>23.46185</td>
<td>-1.880</td>
<td>17</td>
<td>.077</td>
</tr>
<tr>
<td>narrative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>8</td>
<td>64.9250</td>
<td>27.62151</td>
<td>-1.534</td>
<td>17</td>
<td>.600</td>
</tr>
<tr>
<td>female</td>
<td>11</td>
<td>71.1455</td>
<td>23.08490</td>
<td>-1.534</td>
<td>17</td>
<td>.600</td>
</tr>
<tr>
<td>argumentative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>8</td>
<td>53.3875</td>
<td>22.06014</td>
<td>-1.053</td>
<td>17</td>
<td>.958</td>
</tr>
<tr>
<td>female</td>
<td>11</td>
<td>53.7909</td>
<td>10.80560</td>
<td>-1.053</td>
<td>17</td>
<td>.958</td>
</tr>
</tbody>
</table>

Table 4. Paired Samples T-tests on different genres

<table>
<thead>
<tr>
<th>Paired Differences</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
<th>t</th>
<th>df</th>
<th>Sig (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive-Narrative</td>
<td>-12.18</td>
<td>30.63114</td>
<td>7.02727</td>
<td>-1.734</td>
<td>18</td>
<td>.100</td>
</tr>
<tr>
<td>Descriptive-Argumentative</td>
<td>2.72</td>
<td>23.84113</td>
<td>5.46953</td>
<td>.97</td>
<td>18</td>
<td>.625</td>
</tr>
<tr>
<td>Narrative-Argumentative</td>
<td>14.90</td>
<td>22.07843</td>
<td>5.06514</td>
<td>2.943</td>
<td>18</td>
<td>.009</td>
</tr>
</tbody>
</table>
Figure 1. Difference between precision and significance
(Taken from the help section of MATLAB software)

Figure 2. Main structure of the system

Rule 1: If MI is poor and IU is poor then F score is very low.
Rule 2: If MI is poor and IU is average then F score is low.
Rule 3: If MI is poor and IU is good then F score is average.
Rule 4: If MI is average and IU is poor then F score is low.
Rule 5: If MI is average and IU is average then F score is average.
Rule 6: If MI is average and IU is good then F score is high.
Rule 7: If MI is good and IU is poor then F score is average.
Rule 8: If MI is good and IU is average then F score is high.
Rule 9: If MI is good and IU is good then F score is very high.

The inputs are crisp (non-fuzzy) numbers limited to a specific range.
All rules are evaluated in parallel using fuzzy reasoning.
The results of the rules are combined and distilled (defuzzified).
The result is a crisp (non-fuzzy) number.
Figure 3. The membership functions

AND
Figure 4. Fscorer window

Figure 5. Line graph for different genres