

Predicting from Introductory Paragraphs: A Pilot Study in Holistic Assessment of College Level Academic (Remedial) Writing

Swathi Vanniarajan
(San José State University)

Abstract

A study of the relation between the quality of introductory paragraphs of an essay and the overall quality of an essay is important for various reasons. Primarily, such a study can provide us with information about how consistent the budding student writers are in the art of writing academic essays. Secondly, such a study can also provide us with information about the raters' perceptive process itself. Does their initial perception of the essay (i.e., how well or badly the introductory paragraph is written) affect their overall impression of the essay? The results of this study, in overall, indicated that raters' perceptive process does indeed change during the intra-essay reading process and that learners are inconsistent in their writing process with mismatch between the quality of the introductory paragraphs and the subsequent parts of essay they write leading to marginality in the holistic assessment process.

Introduction

An attempt was made in this research study to explore whether it would be possible for two experienced reader-raters to infer whether an essay would pass or fail based on their evaluation of the first two paragraphs (hereafter introductory paragraphs) of an essay alone. The results of the study indicated that even for experienced readers who have been familiar with student writings and holistic assessment for more than 15 years, it would not be possible to predict whether or not an essay would pass or fail based on an holistic evaluation of the introductory paragraphs of an essay alone, due to the inconsistent nature of student writings.

The reason for choosing introductory paragraphs for this study was that the organization and content of an academic writing are predictable, such that the quality of the first two paragraphs can be generalized to the whole essay. In general, the first two paragraphs of an academic essay contain an introduction and a paragraph that is part of the main body of the text. A well written introduction generally informs the reader-rater that the student-writer has interpreted the prompt accurately, synthesized the important ideas, reflected on these, decided on his/her own position, and worked out the organizational

coherence of the essay. A well-written paragraph following the introduction provides further evidence to the reader-rater that the writer can indeed handle the topic adequately. Further reading of the student writing will only be to confirm whether or not the writer has fulfilled the expectations raised in the first two paragraphs of the essay.

The reason for choosing experienced reader-raters was that experienced reader-raters are generally very familiar with student writings; with experience in teaching and assessing writing comes an internalized knowledge of how student essays are generally written even in stressed examination settings. It was, therefore, assumed that experienced-raters would be in a position to speculate (predict) whether or not an essay would pass based on their reading of its introductory paragraphs alone.

Speculation (prediction) in reading and holistic assessment: An unavoidable natural event

Information processing research, a branch of cognitive psychology, identifies reading as an interactive process: an interaction between bottom-up processing and top-down processing. Bottom-up processing refers to the process of using the information received by the sense organs. Top-down processing, on the other hand, refers to the process of using related information that is already stored in one's long term memory in the form of schema knowledge. Information processing research (Cohen, Eysenck, & Le Voi, 1986, Rumelhart, 1977) indicates that reading cannot be based solely on bottom-up processing as it would entail a slow and tedious decoding process. In order to accelerate as well as complete the process of interpretation in real time, top-down processing influences the decoding process by making readers predict what the form and content are likely to be. Research in information processing approaches to reading (McClelland 1987; Rumelhart, 1975; Shiffrin & Schneider, 1977; Ruddell & Speaker, 1985) suggests that activation of top-down processing is not only a natural event in the reading process, but also an unavoidable consequence of acquiring fluency in it. According to Smith (1975), top-down processing can be defined as the prior elimination of unlikely alternatives and that it speeds up the reading process by helping to overcome the limitations imposed by one's working memory¹. Simply put, bottom-up processing is the process of identifying what an item is whereas top-down processing is the process of figuring out what it is likely to be. These two processes coupled together constitute the process of reading.

Carrell (1988) proposes that skilled readers constantly shift between top-down processing and bottom-up processing depending on the demands made by the nature of the reading tasks on the internal

processing resources. She adds that skilled readers are those who are able to monitor their developing interpretation of the text by continually checking it against incoming stimuli. Therefore, if the interpretation (top-down processing) is not confirmed by the subsequent stimuli, then they reject that interpretation and opt for one that is consistent with current as well as previously processed information². She adds that skilled readers constantly shift between top-down processing and bottom-up processing depending on the complexity of the subject matter, linguistic stimuli, and their knowledge of the subject matter. The general view is that dependence on one or the other will lead to marginality in meaning comprehension since dependence on bottom-up processing will result in what is called text-boundedness and over reliance on top-down processing will result in what is called schema interference (Rumelhart 1977, Samuels and Kamil 1984).

Tsui and Fullilove (1998) studied how over reliance on top-down processing could affect the comprehension process in reading. For the purposes of their study, they identified two schema types: matching schemas and non-matching schemas. The findings indicated that in the former, or that when there was a match between top-down processing and bottom-up processing, speeded up comprehension of the reading materials occurred in all readers; on the other hand, in the other, or that when there was a mismatch between top-down processing and bottom-up processing, for those readers who relied more on top-down processing, the real meaning of the text and the meaning of the text as provided by the subjects differed.

The most important point is that if reading is an interactive process, then holistic assessment of writing is also an interactive process since what the reader-rater does during the holistic assessment process is nothing but reading. The difference between the two is that while during the normal reading process one reads for meaning alone, in the holistic assessment process, the reader-raters have to do an additional task, that of attempting to evaluate the student writing against a scoring grid. If top-down processing is a natural and unavoidable event in the reading process, then it is possible that top-down processing occurs in the holistic assessment process also. Needless to say that top-down processing in the holistic assessment process would include not only an expectation of what the form and the content are going to be, but also a speculation as to (prediction about) whether or not the academic writing under evaluation is likely to meet the expected standards. Given this interpretation of holistic assessment process

and the findings by Tsui & Fullilove (1998), it is easy to recognize how the initial impressions made on the raters can significantly affect the scoring process in holistic evaluation is self-evident.

Research Study

Academic English classes in the California State's Oldest University

The university where this study was conducted is one of the oldest universities in the California State University System. The university offers basic remedial English writing classes (called Academic English classes) to those entering freshmen who do not attain the required cut-off score in the English Proficiency Test. The curriculum of Academic English classes focuses on teaching students how to write a 5 paragraph academic essay that is argumentative and persuasive. At the end of the semester, the students in all Academic English classes take a common final exam in which they are expected to reflect for 30 minutes on a controversial prompt and then write a 5 paragraph essay in reaction to the prompt for an hour. The student writings are holistically graded by a pool of teachers who instruct the academic English classes. The holistic assessment sessions begin with an hour long norming session. Following this, each student writing is holistically and anonymously graded by two instructors using a 1 to 6 scale (Appendix 1). Those students who earn an average score of 4 or above are qualified to move on to English 1 A, the second level of the writing classes in the university.

Research Methodology

Two reader-raters who had been teaching the Academic English classes for more than 15 years and had taken part in the holistic assessment every semester ever since they started teaching these classes in the university were recruited for the study. For the purposes of this research study, 100 student essays, 50 of which were awarded a passing score of 4 or above and 50 of which were awarded a failing score of 3 or below by each of these two reader-raters during the common finals in Spring 2000 were randomly chosen for the study. The essays were written in reaction to a prompt entitled "Parents Need to be Graded Too" .

The student writings were then word processed into a computer, as they were originally written, with all the grammatical and spelling errors intact, using Microsoft WORD. Later, only the first two paragraphs of each of these essays were retained in a separate file. It was made sure that each page in this file contained only the first two paragraphs of an essay. In other words, for each of the 100 essays, a separate page containing the first two paragraphs of an essay was made. It was also made sure that the

entries were made in a random sequence so that it could not be inferred from the sequence whether or not the essay containing these paragraphs was awarded a passing or a failing score during the Spring 2000 common finals. Each of these pages was then numbered from 1 to 100 and two boxes, one containing the word 'pass' and the other containing the word 'fail,' were provided at the bottom of each page. Two sets of these introductory paragraphs were then printed out for evaluation.

On a specific day in Spring 2003, both the reader-raters were asked to come to a common place. Once they were well settled, they were instructed to read the prompt "Parents Need to be Graded Too," and then there was a short discussion among the reader-raters and the researcher for about 25 minutes about what an adequate and appropriate kind of introductory paragraphs would be for an essay that would be written in response to the prompt. Once a consensus was reached, the readers were then instructed that their task was to individually evaluate each of the 100 introductory paragraphs in the set without any consultation between them and to come up with a pass or fail decision for the whole essay. In essence, the two reader-raters were clearly informed that they would be reading only the introductory paragraphs (and would not have access to the rest of the essay) but would have to come up with a passing or a failing score for the whole essay based on the reading of the introductory paragraphs alone. The two experienced reader-raters found the task of holistically evaluating the introductory paragraphs and making predictions based on their quality for the whole essays interesting and willingly took part in the study. It is to be noted here that at no point during the entire evaluation process and the research proceedings were these reader-raters informed that their scores based on the reading of the introductory paragraphs alone would be compared to the scores that they had originally awarded to the corresponding whole essays during the holistic reading of the final exams in Spring 2000.

Data for the study

The data for the study consist of the following:

1. The rater 1 pass/fail scores for each of the 100 introductory paragraphs (these scores will hereafter be referred to as rater 1 scores)
2. The rater 2 pass/fail scores for each of the 100 introductory paragraphs (these scores will hereafter be referred to as rater 2 scores)

3. The pass/fail scores as awarded by these two raters to each of the 100 essays containing these introductory paragraphs during the holistic assessment reading session held at the end of the Spring 2000 semester (These scores will hereafter be referred to as Spring 2000 scores).

Hypothesis

The hypothesis for the study can be stated as follows:

Experienced reader-raters, due to their experience in holistic assessment and familiarity with student writings, will be able to predict whether or not a student essay will pass or fail (if not the exact score) based solely on their reading of the introductory paragraphs of the essay alone. In other words, there will be a positive relationship between the quality of introductory paragraphs and the quality of the corresponding essays containing them as perceived by the reader-raters during their holistic evaluation process of the individual essays.

Findings

Table 1 contains findings on the raters' pass/fail scores for introductory paragraphs under study:

Table 1: Raters' scores based on the assessment of introductory paragraphs

	Pass	Fail	Total
Reader 1	35 (35%)	65 (65%)	100 (100%)
Reader 2	58 (52%)	42 (42%)	100 (100%)

It can be seen from Table 1 that Reader 1 had awarded 35 passing scores and 65 failing scores while reader 2 had awarded 58 passing scores and 42 failing scores based on their assessment of the introductory paragraphs under study.

Further analyses indicated that rater 1 had a success rate of 65% in predicting whether an essay would earn a passing or a failing score based on her assessment of introductory paragraphs alone. However, there was a difference in the manner her passing and failing scores for the introductory paragraphs agreed with the original scores. She had predicted that 35 essays would earn passing scores based on her assessment of introductory paragraphs; out of these 35, only 25 had actually passed as per the Spring 2000 scores. This translates to a success rate of 71% in predicting passing scores on the basis of assessing introductory paragraphs alone. She had also predicted that 65 essays would earn failing scores based on her assessment of introductory paragraphs; out of these 65, only 40 had actually failed as per the Spring 2000 scores. This translates to a success rate of 61% in predicting failing scores on the basis of introductory

paragraphs alone. The Wilcoxon test was carried out to determine whether or not the two scores, the original Spring 2000 scores and the scores based on introductions as awarded by rater 1 were significantly different from each other. The test showed that there was a significant difference between the two scores ($Z=-2.535$, $p < .05$). A paired samples t-test was also calculated to confirm this finding. The t-test compared the mean original score to the mean rater 1 score. A significant difference between the two scores was found ($t(99)=2.608$, $p < .05$). With regard to rater 2, analyses indicated that she had a success rate of 62% in predicting whether an essay would earn a passing score or a failing score based on her assessment of introductory paragraphs alone. There was a difference in the manner these passing and failing scores also tallied with the original scores. She had predicted that 58 essays would earn passing scores based on her assessment of introductory paragraphs alone; out of these 58 scores, 35 actually tallied with the original Spring 2000 scores translating to a success rate of 61%. She had also predicted that 42 essays would fail on the basis of her evaluation of the introductory paragraphs; out of these 42 scores, only 27 tallied with the Spring 2000 scores translating to a success rate of 64%. The Wilcoxon test that was carried out to determine whether or not the two scores, the original Spring 2000 scores, and the scores awarded by rater 2 were significantly different from each other. The test showed that there was not a significant difference between the two scores ($Z=-1.298$, $p > .05$). A paired samples t-test was also calculated to confirm this finding. The t-test compared the mean original score to the mean rater 1 score. No significant difference between the two scores was found ($t(99)=-1.302$, $p > .05$).

The Wilcoxon test was carried out to determine whether or not the scores awarded by rater 1 and rater 2 were significantly different from each other. The test showed that there was a significant difference between the two scores ($Z=-3.781$, $p < .01$). A paired samples t-test was also calculated to confirm this finding. The t-test compared the mean rater 1 score to the mean rater 2 score. A significant difference between the two scores was found ($t(99)=-4.064$, $p < .01$).

In order to see whether Rater 1 scores, Rater 2 scores, and the Spring 2000 scores were measuring the same construct, Cronbach's alpha was calculated. The alpha obtained was .5554, an indication that the scorers were measuring the same construct only to a low degree. Correlations were calculated to find out whether the finding with regard to alpha could be confirmed. Low correlations between the scores indeed indirectly confirmed that the scores were not associated with each other in a significant manner as can be seen from Table 2.

Table 2: Correlation Matrix

	Original	Rater 1	Rater 2
Original	1.000		
Rater 1	.3145	1.000	
Rater 2	.2431	.3271	1.000

In order to see, how these differ from one another, a tally analysis of agreement among rater 1, rater 2, and original Spring 2000 scores furthermore showed that there were 8 different combinations of agreement and disagreement among these three scores. The findings with regard to this analysis is given in Table 3.

Table 3: A tally analysis of score agreement among rater 1 (introductions), rater 2 (introductions), and the original scores (for the whole essay).

Score source	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7	Type 8	Total
Rater 1	Fail	Pass	Fail	Pass	Fail	Fail	Pass	Pass	Pass =35 Fail =65
Rater 2	Fail	Pass	Fail	Pass	Pass	Pass	Fail	Fail	Pass =58 Fail =42
Original score	Fail	Pass	Pass	Fail	Fail	Pass	Pass	Fail	Pass =50 Fail =50
Total	24	21	11	7	16	14	4	3	100

Table 3 shows that there were 8 different types of combinations of agreement and disagreement among the three scores; the information can be summarized as follows:

1. Agreement between rater 1, rater 2, and Spring 2000 scores (Types 1 & 2): 45 cases
2. Agreement between rater 1 and rater 2 only (Types 3 & 4): 18 cases
3. Disagreement between rater 1 and rater 2 (Types 5, 6, 7, & 8) 37 cases
 - 3a. Agreement between rater 1 and Spring 2000 scores (Types 5 & 7): 20 cases
 - 3b. Agreement between rater 2 and Spring 2000 scores (Types 6 & 8): 17 cases

In overall, the findings of the study can be summarized as follows:

- ! Rater 1 scores and Spring 2000 scores were significantly different from each other.
- ! Rater 2 scores and Spring 2000 scores were not significantly different from each other.
- ! Rater 1 scores and Rater 2 scores were significantly different from each other.

- ! Rater 1 scores, Rater 2 scores, and Spring 2000 scores were significantly different from one another.
- ! There were 8 different types of combinations of agreement and disagreement among the Rater 1 scores, Rater 2 scores, and Spring 2000 scores.

The results, in overall, indicate that for rater 1, the hypothesis that there is a positive relationship between the perception of the quality of introductory paragraphs and the perception of the quality of the corresponding essays containing them," cannot be supported and has to be rejected on the basis of the results obtained. The results, however, indicate that the hypothesis does have a support in the case of Rater 2, since there were no significant differences between her scores and the Spring 2000 scores. Yet, note that the correlation coefficients between the three sets of scores including the Cronbach's alpha were only to a low degree.

Discussion

Based on the numbers given in Table 3, it was possible to divide the writing samples of students into three types: (i) essays that were consistently good (good introductions and good essays), (ii) essays that were consistently wanting (wanting at the level of introductions and in overall), and (iii) essays that were inconsistent (with a mismatch between the quality of introductory paragraphs and the subsequent parts of the essay).

Essays that were consistently good

Rater 1 had predicted that 35 essays would earn passing scores and rater 2, 58 essays. Table 3 indicated that for 21 essays, there was agreement among rater 1, rater 2, and the Spring 2000 scores. These essays represented a typical passing essay. Productwise, these could be considered to be consistently good from the introduction to the conclusion. The writers of these essays could be said to have mastered the skill of academic writing as shown by the data as their essays earned passing scores at both the introductory and the whole levels of holistic assessment.

Essays that were consistently wanting

Rater 1 had awarded 65 introductions failing scores and rater 2, 42 essays. Table 3 shows that for 24 essays, there was agreement among rater 1, rater 2, and the Spring 2000 scores. Productwise, these could be considered to be consistently wanting from the introduction to the conclusion since evaluation at both the introductory and the whole level indicated that they were indeed failing essays. Such essays generally lack qualities that are desirable in an academic essay. Since these essays earned failing scores at both the

introductory and the whole levels of holistic assessment, it can be claimed that these were the representative samples of essays that were consistently ineffective and defective right from the start. In other words, these essays represented a typical failing essay.

Essays that were inconsistent

In the present context, essays that were inconsistent can be divided into two types: those which contained (1) introductory paragraphs whose (speculative) scores were not in agreement with the scores given for the whole essays and (2) those which contained introductory paragraphs that earned split scores from its two scorers. While the former can be considered to be consisting of those essays that suffered from a lack of internal consistency, the latter can be considered to be containing those essays that suffered from a lack of inter-rater reliability.

Consider Type 3 and Type 4 from Table 3. For Type 3 which contained 11 introductory paragraphs, both raters had awarded failing scores based on their impressions of the introductory paragraphs. The readers' judgements proved to be incorrect because the student writers of these essays, though had written ineffective introductory paragraphs lacking in certain crucial aspects, seemed to have done better in the subsequent parts of their essays as otherwise the essays would not have earned passing scores in Spring 2000 from the same raters. Therefore, it would not be illogical to conclude that both reader-raters had perceived during the Spring 2000 scoring process that the writers had somehow compensated for their inadequate introductory paragraphs with good bodies of text and conclusions.

For Type 4, which contained 7 introductory paragraphs, the two raters had awarded passing scores. Note that the Spring 2000 scores for the full essays containing these introductory paragraphs were exactly the opposite. It is possible that though the student writers of the above introductory paragraphs had written adequate introductory paragraphs, they had not followed them up with effective writing in the subsequent parts of the essays. The fact that they failed to write passable follow-ups to their passing introductory paragraphs suggests that either they had problems finding suitable arguments related to the thesis, or they had difficulty internalizing the process of writing the subsequent parts of an academic essay, or they simply ran out of time. What is important to note here is that both the raters had indeed perceived the writers' failure to match a good body of text with the passing introductory paragraphs during the Spring 2000 scoring process as otherwise the essays containing these introductory paragraphs would not have earned failing scores from both of them.

Type 3 and Type 4 essays, overall, can be said to exhibit that a substantial portion of student writers suffer from inconsistency in their writings; they have mastered the necessary skills for writing certain portions of the essay but not all of the skills for writing all parts of the essay. Of the two types, Type 3 writers of introductory paragraphs seemed to have fared better since they seemed to have overcome their lack of skill in writing good introductory paragraphs and were able to provide good follow-up passages to compensate for it. Type 4 writers, on the other hand, seemed to have fared worse since the subsequent parts of their essay had failed to live up to the quality of the corresponding introductory paragraphs.

Emergence of Type 3 and Type 4 essays is, more importantly, an indication that the readers give attention to all parts of an essay. They make judgements based on the overall impact of an essay, as otherwise the inconsistency in the scores shown in the Type 3 and Type 4 essays (between the speculative scores for introductory paragraphs and the Spring 2000 scores for the overall essays) would not have arisen. It is also interesting to note here that a bad introduction followed by good body of the text works to the student writers' advantage than a good introduction followed by a weak body of the text. Does this mean that the latest impression counts more in the final impression ultimately counts than the initial impression? It is indeed a research question for a follow-up study.

Issues with regard to inter-rater reliability

Table 3 shows that there were 37 cases in which both the raters did not agree with each other. What is noteworthy is that as can be seen from Table 3, of these 37 scores that suffered from split decisions, rater 1 had judged 30 introductory paragraphs (Type 5 and Type 6 together) to obtain failing scores and 7 introductory paragraphs (Type 7 and Type 8 together) to obtain passing scores, while rater 2 had judged exactly the opposite. Note that rater 1's judgement mostly agreed with the failing scores while rater 2's judgement mostly agreed with the passing scores. In other words, it can be said that the assessment of these paragraphs was suffering from "measurement error due to rater errors."

It is possible that rater 1 had higher expectations compared to rater 2 and had preferred to take a punitive approach while rater 2 had less expectations and had preferred to take a more tolerant approach to assessment. Note in this regard that while rater 1 had predicted that 65 essays would fail overall, rater 2 had predicted that only 42 would fail. It is therefore possible to claim that different raters indeed have different expectations in tune with their respective approaches to assessment. In short, the findings indicate that in spite of their vast experience, the expectations of reader-raters may vary and it may show up ultimately in their evaluations. As the findings indicate, it is generally the inconsistent essays that suffer from inter-rater reliability.

Conclusion

The most important finding of this study is that it is not possible to predict the success or the failure of the student essays from a reading of the introductory paragraphs alone. Consistently good or wanting essays are indeed comparatively easy to grade as they may not make heavy demands on one's reading process and are not prone to split decisions from the reader-raters. Consistently inconsistent essays (Types 3 through 8), on the other hand, seemed to be the ones that are susceptible to discrepant scores. The reality is that differences between raters do also exist, and this, in turn, means that borderline essays would always be difficult for the reader-raters to decide on their real worth. What is important in this research study is that the results clearly indicate that rater impressions do change during the intra-essay reading process and that raters do indeed read from the beginning to the end before awarding a score.

The study of introductory paragraphs and its low predictability of overall scores indicates that success in the learning of academic writing as a skill depends upon the successful learning of various phases and their execution. The successful writing of the introductory paragraph with all its desirable characteristics is only one of the phases. As such, the writing process can be deemed to have been internalized only if the student-writer produces a consistent performance at all stages of the writing process. Only when the writer is able to provide evidence that he/she is able to produce a good introductory paragraph, a good body of the text, and a good conclusion, we can make a claim that the writer has internalized all the stages of the writing process. The findings also indicate that the acquisition of various parts and phases of writing an academic essay is not linear at all as otherwise there would not have been essays with well written body of text and conclusions but wanting introductory paragraphs. That the student writers can either compensate for an ineffective introductory paragraphs or compromise well written introductory paragraphs by not following them up effectively in the subsequent parts of the essay is an important finding that is worth noting. Therefore, it may become important for the instructors to give feedback on the various sections of an essay during the in-class essay scoring process. It is suggested here that for the in-class essays, it may be better to use analytical scoring process with feedback on all aspects and phases of the essay while holistic assessment process is used for common finals.

In conclusion, the results indicate that only those essays that are consistently good or wanting are easy to grade without any controversies. Marginal writing is highly subject to many variables, including the raters' approaches to assessment. As such, the prediction of a score for an essay based on the rater's

impression of the introductory paragraph could work only for consistently good or wanting essays and not for marginal ones. Last but more important is the finding that different skills seem to be involved in different parts of essay writing as otherwise a majority of essays could not have been internally inconsistent. The existence of such essays is a reminder that internal inconsistency in student writing and raters' approach to these inconsistencies are some of the important variables that contribute to the degree of confidence in interrater reliability in holistic evaluation.

Limitations

This study has some important limitations. First of all, it should be admitted here that it is difficult to value the introductory paragraphs of essays as a stand alone writing samples and predict the overall effectiveness of an essay since the effect of a writing piece essentially depends on its effect as a whole and not as a combination of separate parts. Writing is an integrative process and not an additive process. Yet, the reader-raters were able to overcome this perceived limitation.

Another limitation of this study is that what had been evaluated could be an unknown variable that has not been studied as yet: a student writer's ability to write effective introductory paragraphs and its relationship to sustained writing. It is important to note here that this may be a totally different skill than the writing of an essay. As such, predicting the overall score of an essay based on an evaluation of this phase of the writing process may turn out to be a false assumption. In spite of these limitations, this study was worth carrying out since it had given insights into why certain student essays are easy to assess and why some are not and how the reader-raters' perceptive process works while holistically grading a student-written essay.

End notes

1. According to Smith, top-down processing is predicting and hypothesis testing and that it is due to brain's sluggish rate of information processing though he actually refers to working memory limitations in information processing.
2. It is important to note here that interactive processing occurs at both sentential and discourse levels. For example, experiments on the interpretation of garden-path sentences have clearly shown that even experienced readers notice the inconsistency between the first and the second part of these sentences and have to rework on their interpretation.

References

- Carrell, P.L. (1988). Some causes of text-boundedness and schema interference in ESL reading. In P.L. Carrell, J. Devine, & D.E. Eskey, (Eds.), *Interactive Approaches to Second Language Reading*. Cambridge Applied Linguistics Series. Pp. 101-113.

Cohen, G., Eysenck, M.W., & M.E. Le Voi. (1986). *Memory : A Cognitive Approach*. Philadelphia, PA.: Open University Press.

Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, 4:126-135.

McClelland, J.L. (1987). The case for interactions in language processing. In M. Coltheart (ED.), *Attention and Performance XII: the Psychology of Reading*. Hove, UK: Lawrence Erlbaum Associates Ltd. Pp. 3-36.

Ruddell, R.B. and R.Speaker. (1985). The interactive reading process: A model. In H.Singer and R.B.Ruddell (Eds.), *Theoretical Models and Processes of Reading*. Newark, Delaware: International Reading Association. Pp. 751-793.

Rumelhart, D.E. (1975). Notes on a schema for stories. In D.G. Bobrow & A.M. Collins (Eds.), *Representation and Understanding: Studies in Cognitive Science*. New York: Academic Press. Pp. 211-236.

Rumelhart, D.E. (1977). *Introduction to Human Information Processing*. Chichester: Wiley.

Samuels, S.J. and M.L. Kamil. (1984). Models of the reading process. In P.D. Pearson (Ed.), *Handbook of Reading Research*. New York: Longman.

Shiffrin, R.M. & W. Snyder. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84:127-190.

Smith, F. (1975). The role of prediction in reading. *Elementary English*, 52 (3): 305-311.

Tsui, A.B. M. & J. Fullilove. (1998). Bottom-up or top-down processing as a discriminator of L2 listening performance. *Applied Linguistics*, 19 (4): 432-451.

Appendix 1

CSU ENGLISH PLACEMENT TEST SCORING GUIDE

Score of 6: Superior

A 6 essay is superior writing, but may have minor flaws.

- § addresses the topic clearly and responds effectively to all aspects of the task
- § explores the issues thoroughly and in depth
- § is coherently organized, with ideas supported by apt reasons and well-chosen examples
- § has an effective, fluent style marked by syntactic variety and a clear command of language
- § is generally free from errors in mechanics, usage, and sentence structure

Score of 5: Strong

A 5 essay demonstrates clear competence in writing. It may have some errors, but they are not serious enough to distract or confuse the reader.

- § clearly addresses the topic, but may respond to some aspects of the task more effectively than others
- § show some depth and complexity of thought
- § is well organized and developed with appropriate reasons and examples
- § displays some syntactic variety and facility in the use of language
- § may have a few errors in mechanics, usage, and sentence structure

Score of 4: Adequate

A 4 essay demonstrates adequate writing. It may have some errors that distract the reader, but they do not significantly obscure meaning.

- § addresses the topic, but may slight some aspects of the task
- § may treat the topic simplistically or repetitively
- § is adequately organized and developed, generally supporting ideas with reasons and examples
- § demonstrates adequate facility with syntax and language
- § may have some errors, but generally demonstrates control of mechanics, usage, and sentence structure

Score of 3: Marginal

A 3 essay demonstrates developing competence, but is flawed in some significant way(s).

- § distorts or neglects aspects of the task
- § lacks focus, or demonstrates confused or simplistic thinking
- § is poorly organized or developed
- § does not provide adequate or appropriate details to support generalizations, or provides details without generalizations
- § has problems with or avoids syntactic variety
- § has an accumulation of errors in mechanics, usage, and sentence structure

Score of 2: Very Weak

A 2 essay is seriously flawed

- § indicates confusion about the topic or neglects important aspects of the task
- § lacks focus and coherence, or often fails to communicate its ideas
- § has very weak organization, or little development
- § provides simplistic generalizations without support
- § has inadequate sentence control and a limited vocabulary
- § is marred by numerous errors in mechanics, usage, and sentence structure

Score of 1: Incompetent

A 1 essay demonstrates fundamental deficiencies in writing skills

- § suggests an inability to comprehend the question or to respond meaningfully to the topic
- § is unfocused, illogical, incoherent, or disorganized
- § is undeveloped
- § provides little or no relevant support
- § has serious and persistent errors in word choice, mechanics, usage, and sentence structure

Non-response essays, those that reject the assignment or fail to address the question, should be given to the table reader. Readers should not penalize ESL writers excessively for slight shifts in idiom, problems with articles and confusion over prepositions, and occasional misuse of verb tense and verb forms, so long as such features do not obscure meaning.