

The Relationship between Topic Difficulty and Mode of Discourse: An In-depth Study of EFL Writers' Production, Recognition, and Attitude

Majid Nemati
University of Tehran

Abstract

It has been known for a long time that L1 writers perform differently in different modes of discourse. Despite its importance, there has been little conclusive evidence to shed some light on the issue of ESL/EFL learners' writing performance across various discourse types. Therefore, this research study was designed to investigate differences resulting as the effect of four discourse modes, i.e. Narrative, Description, Explanation, and Argument, on EFL writing skill. To do so, in three different phases, the probable differences were evaluated in three dimensions of production (through eliciting compositions), recognition (through cloze tests derived from compositions written in different discourse modes), and finally the learners and teachers' attitude towards these types of writing (through questionnaires). Results show a statistically significant difference among these discourse types, ranking descriptive and narrative writing as the easiest and argumentative as the most demanding type of writing. The findings of this study confirm that the L2 learners follow the same pattern that has been established for the L1 writers. In other words, argumentative, compared to descriptive or narrative writing is more demanding, a fact which may have substantial consequences on language evaluation.

Keywords: Topic, Discourse mode, Writing, Testing.

INTRODUCTION

Writing has always played a crucial role in students' academic career and this role is currently receiving more momentum. Keller-Cohen and Wolfe (1987) estimated that 70% of courses within the undergraduate curriculum require some sort of writing. Besides, 97% of the faculty surveyed in the same research emphasised the

importance of writing for college study. Therefore, writing is an indispensable component of every reasonable academic syllabus.

Undoubtedly, any form of teaching should be accompanied by a proper method of evaluation. The importance of accuracy of evaluation, especially when it comes to writing, cannot be overestimated, because many important decisions made in higher education, from admission to graduation, are based exclusively on written tests and tasks.

Generally speaking, writing is evaluated in two ways, directly and indirectly. In direct methods of testing writing students are required to create a piece of writing, usually in the form of an essay, whereas indirect methods evaluate writing skill through multiple-choice items, presuming that one's ability to write is manifested in the components of writing such as grammar and vocabulary.

Comparing these two, the direct evaluation of writing skill is generally accepted and practised by most test batteries, even by TOEFL, which has been known for a long time for its enthusiastic adherence to the use of multiple-choice items. However, there are still many controversial issues in the domain of testing writing that, in Raimes' words (1990), are "causes for concern". These concerns, like any other concern, are philosophically due to our lack of enough knowledge. It is admitted by scholars and pioneers of the field (e.g. Hamp-Lyons, 1986) that there is still a lot left to be learned about direct assessment of writing proficiency.

Hamp-Lyons, (1990) distinguishes four major components in a direct test of writing which may affect the test reliability and validity: *the writer, the task, the scoring procedure, and the reader*. These components, each comprising several variables, can significantly change the test results. Among different components of a direct test of writing, the task has been in the focal point of the researchers' attention as the final product and, after all, it is the basis of any judgements and decision makings with regard to the writer's mastery over the writing skill.

Task, itself, is not a single entity, rather a conglomerate of several variables. Each of these variables, in turn, can affect the writer's performance. But, there is one very controversial factor that, even

after the huge bulk of research done on it, still remains challengingly mysterious. This problematic variable is the *topic* or *prompt*.

The works of Brossell and Hoetker Ash (1984), Carlson et al. (1985), and Spaan (1989) indicate that there is a general assumption among the teachers as well as researchers that the nature of writing tasks depends on the given topic and therefore the students' writing scores can change drastically from one topic to another (Hartog, 1936; Hartog et al., 1941; Braddock, Lloyd-Jones, and Schoer, 1963; Britton et al., 1975; Poetker, 1977; Hirsch and Harrington, 1981; Applebee, 1982; Freedman and Calfee, 1983; and Pollitt et al., 1985). However, the counter argument has not been left untouched. Brossell and Hoetker Ash (1984) provided some evidence to show that topic difference is not that much significant. Carlson et al. (1985) found the same correlation coefficients for topics of different type and topics of similar type and Spaan (1989) observed similar performances on two supposedly different topics.

The paradox of these contradictory findings can be accounted for by two factors. First, the outcome of a study depends, to some extent, on the researcher's approach and expectations and as Hamp-Lyons (1990, p.74) put it "the solution one prefers will depend on one's statistics and on the expectations one started with". Secondly, it is vital to realise that topic or prompt is not a single entity (e.g. Ruth and Murphy, 1988) rather it is a complex of many other variables (Hamp-Lyons and Kroll, 1997) such as linguistic complexity, audience, subject matter, rhetorical specification, and finally *mode of discourse* which is the main focus of this study.

Mode of discourse and writing competence

It is still an open question whether or not some writers perform better in certain modes of discourse but not equally well in the others. Modes of discourse have been classified differently in different sources and in some cases one might find them overlapping and confusing. It seems that the lack of a clear-cut definition of different modes has caused at least some of the confusion. Despite this confusing situation, there is unanimity over the names and basic

categorisation of the four primary modes of discourse i.e., *narration*, *description*, *exposition* and *argumentation*. This traditional classification is still overwhelmingly accepted and widely used in present literature (e.g. Scott, 1996). She states the problem quite clearly:

Does writing competence vary with mode of discourse? The four primary modes of discourse traditional to the discussion of writing are *narration*, *description*, *exposition*, and *argumentation*. Standards for a good narrative would differ somewhat from standards for a good description, just as good descriptive writing would differ from good expository or persuasive writing. However, the important question with regard to writing competence involves individual variation in different discourse modes. That is, are some writers competent when writing in one mode but less competent when writing in another? (p.13)

She further maintains that “writers may have competence in several, but not necessarily all, modes of discourse.” (ibid.)

Previous studies

As early as 1953 Kincaid drew attention to the fact that writers, especially better ones, perform differently in different modes of discourse i.e., narrative, descriptive, argumentative, and expository texts. Braddock et al. (1963, p.8) also alarmed that “variations in mode of discourse may have more effect than variations in topic on the quality of writing.” Veal and Tillman (1971) made an observation on second, fourth and sixth grade children’s writings of four different modes of description, argument, explanation and narration and observed that for fourth grade students the scores for expository writing were higher than argumentative modes. Quite contradictory to most research results, Quellmalz and Capell (1979) reported that ratings for narrative essays were lower than the expository ones and suggested three reasons for this phenomenon:

1. There is a tendency among raters to regard expository as superior to narrative.
2. Different discourse modes are not emphasized equally in the curriculum.

3. Students suffer from a lack of knowledge to cope with narrative assignments.

In a frequently cited study, Crowhurst and Piche(1979) observed a variation in syntactic complexity among the texts written by the same writer but in different modes of discourse. During this research project sixth grade and tenth grade children wrote argumentative essays which were syntactically more complex than either descriptive or narrative essays. Previous studies done by Rosen (1969) and then by Perron (1977) had also shown similar results. Crowhurst (1980) studied the writings of six, ten and twelve grade students and reported that narration received higher scores than argument. It was also maintained that there is an association between syntactic complexity and the quality of argumentative writings but that such relation does not exist for narrative essays.

Similarly, Quellmalz, Capell, and Chou (1982) maintain that the level of performance varies from narrative to expository discourse. Freedman and Calfee (1983) who compared compositions requiring quotations and those requiring opinions, argued that even within a certain mode there are sub-modes which require different abilities on the side of the writer. Reed et al. (1985) in an experiment involving English-speaking college students with different writing capabilities found that for these students narration was the easiest, "perhaps to the point of automaticity", description second, and persuasion the most difficult mode of discourse to write in because of different degrees of cognitive capacity engagement each required.

All in all, drawing upon the previous studies Park (1987) makes the following conclusions:

1. Discourse mode is significant with regard to L1 writing performance.
2. Among different modes, argument is the most difficult one and therefore it usually receives the lowest scores.
3. The effects arising from a change in type of discourse is stronger for elementary than for the high school level.

All of the foregoing studies were experimented with L1 subjects. There are only a handful of studies about the situation in L2 of which

few are rigorous enough to be considered as conclusive with regard to any aspect of topic variable, let alone the discourse mode. The study of the effects caused by modification of formality level of language done by Hirokawa and Swales (1986), for example, was too small. The most important and related research study is that of Hamp-Lyons and Prochnow (1994) in which they investigated the effect of level of difficulty of the essay prompt on ESL learner performance. The prompts were divided according to writing task type namely, expository/private; expository/public; argumentative/private; argumentative/public; and a combination of two or more of these types. It was hypothesised that the expository/private was the easiest and the argumentative/public the most difficult one to perform. Surprisingly, they found that the mean writing score increased as topic difficulty increased. In a more recent research study of this type Koda (1993), studying American college students studying Japanese, realised that different linguistic (grammaticality and sentence structure) and rhetorical (organisation and coherence) competencies are required for two different modes of description and narration. And finally, in another study by Hamp-Lyons and Porchnow (1994), using the same categories suggested by themselves (1991) investigated expert judgement of the task difficulty and found out that although these experts shared the same ideas and criteria about what makes a prompt difficult, their predictions were the reverse of the pattern shown by the subjects' actual scores.

It can also be concluded that "there has been, then, no real investigation of the effect of task variable on the measured writing quality of ESL writers on direct tests of writing (Hamp-Lyons, 1990, p.75). Obviously, there is little solid and objective data to pinpoint the topic difference resulting from the mode of discourse and its effect on the ESL writer performance. Therefore, conducting a fully-fledged investigation into this ambiguous but important issue seems to be justifiable and warranted.

The present study: aims and research questions

The present study is an attempt to investigate those aspects of essay topics which may affect the EFL learner's performance due to the

differences in the discourse type of the elicited essay. As discussed earlier, topic is not a single variable, rather it is a conglomerate of variables, each and every of them may affect the writer's performance. The present study mainly focuses on the four basic modes of written discourse i.e. *argument, description, explanation* and *narration* and whether or not EFL writers perform differently while writing in these modes, the nature of essays written in these modes and the attitude of students towards writing in these modes. Accordingly the main three questions of this study can be stated as follows:

1. *Is there any significant difference between the quality of essays written by EFL learners in difficult discourse modes (argument/explanation) with that of those written in relatively easier modes (narration/description)?*
2. *Do the essays written in different discourse types bear different degrees of cognitive and/or linguistic difficulty to comprehend and reconstruct?*
3. *Do the EFL writers prefer certain mode(s) of discourse to other mode(s) while choosing a topic to write on?*

METHODOLOGY

In order to find the answers to these three questions, a three-phase study was designed. The first phase focused on the production of essays in different discourse modes by EFL writers while in phase two the recognition of texts written in these four modes were compared against each other to see if any of them were more difficult to comprehend. The third phase was designed to survey the learners' attitudes towards topics that require a certain type of discourse mode.

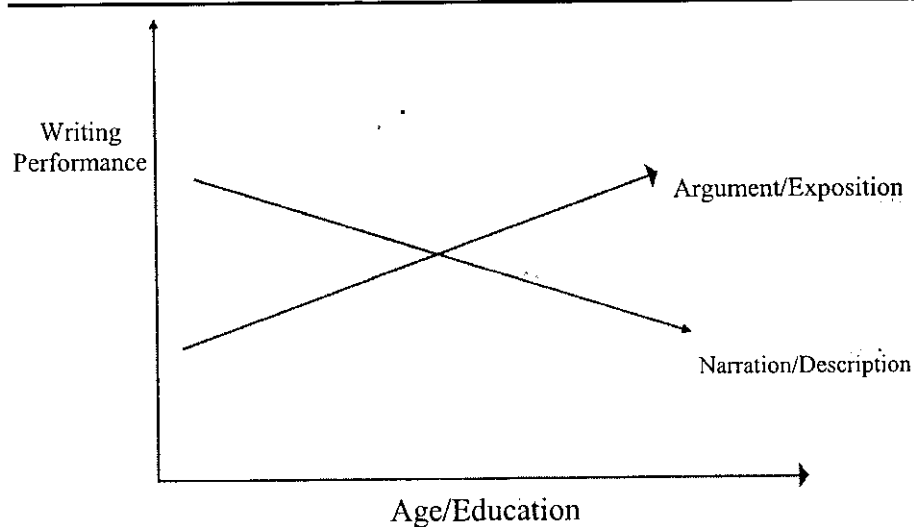
Phase one: production

For differences in production a traditional essay examination is proposed. It was initially intended to compare all four different discourse types against each other in an ideal design in which each subject would have written in all four types of discourse but because of certain practical limitations, it was not plausible to do so.

Unwillingly, each subject, instead, was required to write two compositions, one in a supposedly easy mode (narrative/description) and one in an assumed more difficult mode (argument/explanation). These compositions were later scored and compared against each other.

It has to be pointed out that at the starting point of the study the researcher used to believe that at a tertiary level, argument would not be considered as a difficult mode. This working assumption was in line with that of Park (1987) that the effect of mode of discourse is weaker for high school level than for elementary school level. Therefore, it seemed quite rational to hypothesize that for academic L2 situations at tertiary level where the subjects have been exposed to an enormous number of argumentative texts and have, supposedly, developed a considerable level of critical mind by learning how to argue for or against a point of view, argumentation is not a problem anymore. The following is a schematic view of the researcher's initial assumption:

Figure 1: *The cross-over effect between an easy mode and a difficult mode of writing*



This assumption was actually rejected during the pilot study (Nemati, 1995) in which argument behaved very similarly to the patterns observed in L1 studies. In the next step, then, each member of the former group was paired with one from the latter one: description and argumentation were put into one set, and narration and explanation into another to make two sets of contrasting discourse types.

The main problem in the way of comparing different modes of discourse was to ensure that the probable observed differences were with reasonable confidence attributable to discourse type differences and not to other uncontrolled interfering variables. The easiest solution seemed to be repeating the comparison at least one more time with a different set of prompts. To achieve a high degree of certainty it was decided to repeat the test *four* times, each within a different situation, and aggregate the results.

The final design of the study, after matching the correspondent contrasting compositions, was as follows (Figure 2):

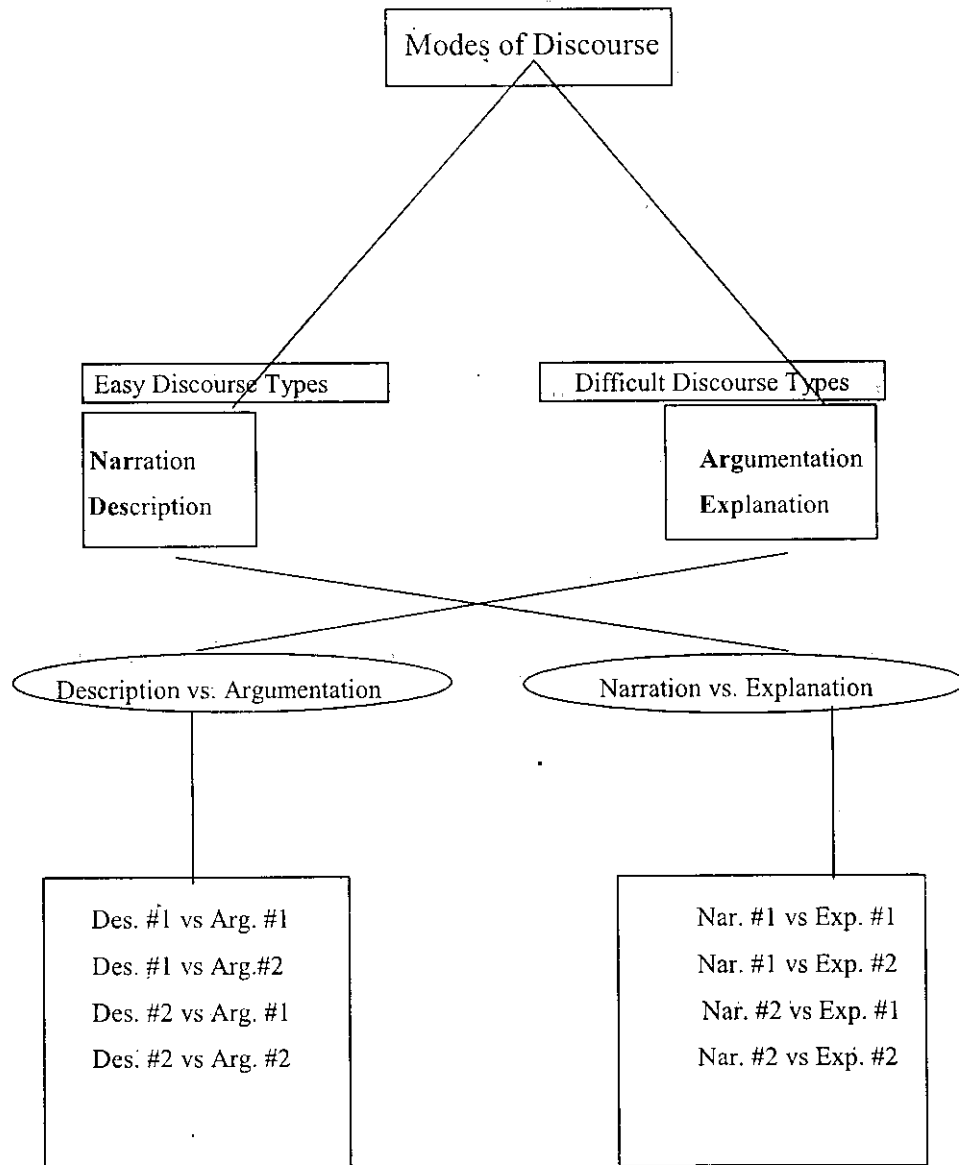
Materials and instrumentation

The process of choosing the prompts was not an easy one. This process is so important that Hamp-Lyons (1988) calls prompts "product before" and so tricky that even a single common word like "discuss" may be interpreted in three different ways (Dudley-Evans, 1988). In fact, a pool of topics was created by gathering topics from four different sources:

- consulting several writing textbooks available
- topics used in other similar research projects e.g. Hamp-Lyons and Prochnow (1991).
- asking EFL students, including Iranian ones, to list topics that they believed they would be likely to see in a language test.
- using the experience and the intuition of language teachers.

Therefore, for each type of discourse two carefully selected essay prompts were suggested. The prompts are given in Appendix 1.

Figure 2: The design of dividing opposing discourse types into sub-group classes taking part in the experiment



Subjects

The subjects of this phase of the study were chosen randomly from English language classes held in various Iranian universities located in Tehran wherever accessible, with no intentional preference whatsoever. Unfortunately it was practically impossible to test the subjects' language proficiency level but, as these students have already passed a very competitive test, the National University Entrance Exam, it can be concluded that the subjects' language proficiency levels ranged from lower intermediate to advanced level for Masters students. This inference is evident from the piece of writing they have produced. However, this variety of command over the English language does not affect the reliability of the research since individuals are compared against themselves; in fact, it rather increases the generalizability of the findings.

The number of subjects in the largest group is as high as 34 and in the smallest group shrinks to 17. Altogether, there are 103 subjects in the **Description versus Argumentation** (Des. vs Arg.) main group and 80 in the **Narration versus Explanation** (Nar. vs Exp.) main group. The total number of subjects taking part in this part of the study is 183 creating 366 compositions to be investigated.

Scoring Procedure

After collecting the compositions, the difficult job of scoring had to be done. After a co-ordinating session, two experienced NNS teachers embarked on the scoring job, judging solely by the impression made by the essays. Besides, they also went through the guidelines suggested by Farhady et al. (1995) for marking ESL compositions on the 20-point scale with decimals of 0.25 which is the common marking scale in Iran. The inter-rater reliability was estimated at 0.92.

Phase two: recognition

This phase of the study was mainly intended to examine the nature of texts produced in different discourse types. The production stage was to investigate how well ESL learners produce a piece of linguistic discourse whereas this part of the research project was developed to

compare and contrast the nature and features of the texts produced in different modes of discourse and measure their cognitive level of difficulty.

Materials and instrumentation

In this phase of the study, four texts, each in a different discourse type, had to be compared. It would be very difficult, if not impossible, to find four parallel texts, equal to each other in every aspect except for the type of discourse. Alternatively, it was decided to use four texts written by the same person, and therefore, originated from the same level of linguistic competence, so that if any difference exists, it should be a genuine one resulting from the different prompts which led to these texts. An Iranian research student was assigned to produce the texts. He was not told about the research project and its underlying assumptions to keep him impartial towards different essay topics. All the aforementioned eight prompts were offered to him in pairs so that he had a choice to pick his favorite topic from each mode. He chose the Argumentative prompt #2, Descriptive prompt #2, Narrative prompt #1 and Expository prompt # 2.

The written texts, then, were turned into standard cloze tests. In other words, the very first and last sentences of each passage were left intact and in all other sentences every 7th word was deleted. In order to limit the choice of words fitting correctly into the blanks the first letter of each missing word was provided. This modified version of cloze test is similar to the one developed by Davies (1965) in which he provided only the first letter of each missing word.

From each of the four essays a passage comprising 15 blanks was chosen. Therefore, the final test had 60 missing words to be filled within 40 minutes. The test administration supervisors divided the allotted time into four equal intervals of 10 minutes and made the students move to the next passage at the end of each ten minutes. This way, attempt was made to achieve an even distribution of time and effort among the four passages.

Subjects

The subjects of the Recognition phase were chosen from three universities located in Tehran: Tehran University (Faculty of Foreign Languages), Open University (Northern branch), and Open University (Southern Branch). The profile of the subjects is given in the Table 1.

Table 1: The profile of subjects taking part in the recognition phase of the study

University	Number (N)	sex		level	
		(M)	(F)	(Under)	(Post)
Tehran University	40	25	15	32	8
Open University (North)	26	0	26	26	0
Open University (South)	34	34	0	34	0
Total	100	59	41	92	8

Scoring procedure

There are mainly two methods of scoring cloze tests: the *exact word method* and the *acceptable word method*. In exact word method the testee gets the point only when the filled word is the same as the word written in the original text. In the other method, *the acceptable word method*, all the words that can fill the gap correctly will receive the point. The latter method was chosen for scoring the cloze tests.

Subjects, after the scoring procedure, received four sets of scores, each representing their performances in reconstructing written language of a certain mode of discourse. A copy of the cloze test used in this study is provided in Appendix 2.

Phase three: preference and attitude

One of the areas in the field of evaluating writing which has drawn some attention towards itself is the issue of how the learner writers choose a topic. Despite this attention, it seems that "previous quantitative attempts to examine factors affecting students' choice have been inconclusive" (Polio and Glew, 1996, p.35). They further hold that one of the reasons that may have hindered researchers in finding the factors behind the prompts that may make some topics

more appealing over others is that "most of the research on the writing process examines, and not appropriately, the overall writing process from beginning to the end" (ibid). So the very first step in writing which is choosing a topic has not been studied properly on its own.

An outstanding qualitative study, as already cited, is done by Polio and Glew (1996). They concluded that several factors affect the students' choice of topic including "their own background knowledge, question type, and specificity of the topic" (p.35). Although they did not focus on the effects of the discourse mode, it was observed that personal narrative prompt was chosen because the subject could write immediately and faster as it does not require the writer to plan the essay before starting the actual writing. This is yet another signal that mode of discourse can make a difference in learners' choice of topic and, presumably, in the quality of their writing.

Another instance of such studies is that of Chiste and O'Shea (1988) in which they found that ESL students when offered a choice of topics, unlike native speakers, tend to choose shorter topics containing fewer words. However, as the shorter topics were placed at the top of the list it makes it look possible that the subjects were influenced by the order of the prompts in the list, a pitfall which has been avoided in the present study.

Probably the most frequently cited research of this type is that of Hamp-Lyons and Prochnow (1994). They asked two trained MELAB composition readers and two ESL writing experts to rate 64 MELAB topics (administered in the period 1985-1989) on a three-point scale: easy, average difficulty, and hard. In the next step, using the categories developed by Bridgeman and Carlson (1983) they classified the topics into five task type categories:

1. expository/private
2. expository/public
3. argumentative/private
4. argumentative/public
5. combination

According to the experts' judgments the argumentative/public was considered as the most difficult type of topic and the expository/private as the least difficult type. However, when they tried

to find the probable existing relations between the scores and the rating and the topic type they realized that “mean writing score increased rather than decreased, as topic difficulty increased, except for topics in the group judged as most difficult” (p.62).

Questionnaire design and data collection

The data for this part of the study have been collected through a questionnaire. In this questionnaire a list of all eight composition prompts, used in the production phase, was offered to Iranian university students to rank them according their preference on a five-point Likert scale. The students were supposed to read the topics and, according to the degree that they would prefer to write on those topics in a writing test, rate them from one to five, one signifying the least and five signifying the most favorite topic. The eight topics were arranged in a certain way to avoid the pitfall that is associated, as mentioned earlier, with the study done by Chiste and O’Shea (1988). In this questionnaire the topics were listed in the following order:

1. Argumentative topic #1
2. Descriptive topic #1
3. Narrative topic #1
4. Explanatory topic #1
5. Explanatory topic #2
6. Narrative topic #2
7. Descriptive topic #2
8. Argumentative topic #2

As shown in the above list, there are two topics for each mode and the order for the second set of the topics (from explanatory to argumentative) is the reverse of the first one (from Argumentative to Explanatory).

Subjects

The subjects of this phase of the study were chosen from the Iranian students studying in different universities. Eighty-seven questionnaires have been collected and analyzed. The result of this analysis is presented in the next section.

RESULTS AND DISCUSSIONS

In the previous section, for the convenience of readers, the research study was divided into three phases namely: *production, recognition, and preference*. Similarly, the results are presented in three corresponding sections.

Production

In the production phase of the study, two comparisons were made: argument versus description, and narrative versus explanation. Each of the two groups was actually compared in four separate sub-groups with four different sets of topics and the results were aggregated. And here are the descriptive statistics for argument versus description after being aggregated:

Table 2: Descriptive statistics for the aggregated group of Argument vs. Description

	Mean	SD	Minimum	Maximum	N
Argument	10.60	4.77	3.00	19.50	103
Description	11.04	4.57	3.25	19.75	103

A related t-test was run to investigate any possible statistical differences between the two types of compositions (i.e. description and argument) to answer the first research question. The result is presented in Table 3:

Table 3: The results of related t-test for Argument vs. Description

Mean	SD	SE of Mean	t-value	df	2-tail Sig
-.4381	1.551	.157	-2.78	96	.006**
95% CI (-.751, -.126)					

As shown in the table there is a highly significant difference ($P < 0.01^{**}$) between argument and description essay scores written by the same subjects. Therefore, the first null hypothesis is rejected and the experimental hypothesis is accepted. The conclusion is that descriptive topics, compared to argumentative ones, seem to be much easier for EFL learners to write on.

A similar procedure is given below to illustrate the observed relationship between narrative and explanation.

Table 4: Descriptive statistics for the aggregated group of Explanation vs. Narrative

	Mean	SD	Minimum	Maximum	N
Narrative	7.67	3.08	2.75	13.75	80
Explanation	7.97	3.01	2.00	13.75	80

Based on the rationale stated for the application of matched t-test for Argument/Description group, a similar test has been run for Narration/exposition group, too. The result of the test is given in Table 5.12.

Table 5: The results of related t-test for Explanation vs. Narrative

Paired Differences					
Mean	SD	SE of Mean	t-value	df	2-tail Sig
.2908	1.520	.170	1.71	79	.091
95% CI (-.047, .629)					

Although there is a trend here in the direction which would be needed to reject the null hypothesis, the results do not indicate a difference large enough to reject the second null hypothesis. It can be concluded that narratives and explanatory compositions do not meaningfully differ from each other at the production level.

Discussion

The statistical analyses have indicated a difference between argument and description but not for narrative and exposition. It might be reassuring that a more robust t-test was also run which confirmed the present results. The first interesting point is that argument quite obviously stands out as a much more difficult mode of discourse even in a foreign language situation. This is quite contradictory to the researcher's prediction, discussed earlier. It can be concluded that writing an argument is quite demanding even for critically-developed minds of students at tertiary level.

Considering the lack of a significant difference between narration and exposition, two tentative reasons jump to the mind. First, it seems

that exposition is a much too broad category. It can include a range of rhetorical features which may have different bearing on the difficulty level of the topic. A similar research study, comparing different varieties of exposition, may shed some light on the issue. Second, one of the narrative topics proved to be culture-bound and hence not quite familiar for Iranian students. This made narration look a bit more difficult than it could be. The selection of this topic was actually an over-generalization of Anglo-American academic culture into Iranian academic situation. Topics like "the biography of a person who rose from a humble background" are typical Western composition subjects. If this is the case, then it can be concluded that other variables, such as subject matter familiarity can override the effect of mode of discourse. This is an alarming finding for topic developers at major test batteries or teachers of English for academic purposes.

Recognition

The second part of the study is mainly concerned with the nature of the texts produced across the four main modes of discourse. As mentioned before, four different texts written by the same person under similar conditions were turned into modified cloze tests. These tests were later given to a number of Iranian students. The following is the result observed:

Table 6: Descriptive statistics for the four different sections of the cloze test

	Mean	SD	Minimum	Maximum	N
Cloze (Exp)	8.32	2.68	1.00	14.00	100
Cloze (Nar)	11.87	1.87	5.00	15.00	100
Cloze (Des)	7.49	2.82	1.00	14.00	100
Cloze (Arg)	6.75	2.68	1.00	13.00	100

A repeated-measure ANOVA was run to check any significant differences. The result is presented in Table 7.

Table 7: Analysis of variance for the variable Cloze Test by variable MODE

Source of Variation	SS	DF	MS	F	Sig of F
WITHIN+RESIDUAL	841.18	297	2.83		
TEXT TYPE	1542.57	3	514.19	181.55	.000

As the table shows all four groups are significantly different from each other. Once again, argumentative mode of writing, with a mean score of 6.7 proved to be the most difficult text, while narrative, compared to other discourse types, achieved a relatively high mean score of 11.8 and therefore is, by far, the easiest.

The next step is to determine where the difference lies, or to put it in a simpler way, which two of these modes are significantly different from each other. The most common way of achieving this is to compare each group with every other through a multiple comparison test. Among the available multiple comparisons tests (e.g. the Tukey and the Sheffe tests) the Tukey test is "possibly more sensitive . . . [and] does tend to give a more accurate Type I experiment wise error rate" (Youngman, 1979, p.84). Youngman (ibid) also strongly stresses the need for a sample of at least 10 members. This test is usually intended for use with equal-sized groups. Considering these conditions and recommendations, a Tukey test with a significance level of $p < 0.05$ was run to locate the existing differences among the groups. Table 8 presents the results.

Table 8: The Tukey Test for different types of texts used in the cloze

Mean	MODE	Grp 4	Grp 3	Grp 1	Grp 2
6.7500	Grp 4				
7.4900	Grp 3	*			
8.3200	Grp 1	*	*		
11.8700	Grp 2	*	*	*	

(*) Indicates significant differences

Group1= Exposition, Group2= Narration, Group3= Description, Group4= Argumentation

As Table 8 shows, all four groups are significantly different from each other. Once again, the argumentative mode of writing, with a mean score of 6.7 proved to be the most difficult text, while narrative writing, compared to other discourse types, achieved a relatively high mean score of 11.8 and therefore is, by far, the easiest.

In a further peripheral investigation to make sure that texts written by the same person in different discourse types are of different nature and possess different levels of difficulty, all four passages previously used as cloze tests were examined to determine their readability estimate. The next table presents the results.

Table 9: Readability estimates for different discourse types used in the cloze

	Argument	Description	Narrative	Explanation
Flesch Reading Ease	46.0	60.8	77.9	55.3
Flesch-Kincaid Grade Level	14.0	12.2	6.7	10.4
Coleman-Liau Grade Level	10.8	9.3	8.0	12.0
Bormuth Grade Level	10.3	10.0	9.5	10.6

In this analysis all four formulas, except Flesch-Kincaid which considers Description slightly more difficult than Explanation, consistently ranked argument as the most difficult text, then expository, then description and finally narration, for their readability level.

Discussion

The findings of this phase remind us that the texts produced in different modes of discourse would be of different nature. This prevents us to falsely conclude that argument is difficult because students are not trained to write arguments, or narration is easier simply because they have been exposed to it from early stages of schooling. As the results show different modes of discourse have different cognitive demands on the mind of the reader and these texts are different by nature and hence creating them is not a similar task.

Preferences and attitudes

The third section of this research study was designed to investigate the EFL learners' preferences toward different topic types. So far it has been shown that the mode of discourse does have a bearing on the writer's performance. Now let's have a look at students' interpretations. Do they really know, consciously or sub-consciously, that writing on a certain topic is more demanding and therefore more

unlikely to achieve a good mark than another one? And if so, how do they rank these essay prompts? Is there any relationship between their ranking and the type of discourse to which the prompts belong?

In a questionnaire the eight topics used to elicit compositions in the first phase of the study were randomly listed and the students were asked to rank them on a Likert-type scale according to their preferences. The statistics for this questionnaire is given in Table 10.

Table 10: Descriptive statistics for the eight topics from four discourse types

	Mean	SD	Minimum	Maximum	N
Arg.#1	2.79	1.29	1	5	87
Arg.#2	2.99	1.42	1	5	87
Nar.#2	3.29	1.36	1	5	87
Exp.#2	3.33	1.24	1	5	87
Exp.#1	3.40	1.31	1	5	87
Des.#1	3.51	1.39	1	5	87
Des.#2	3.62	1.27	1	5	87
Nar.#1	3.80	1.24	1	5	87

It is evident that argumentative topics were the least popular topics among the EFL learners. However, each pair of these eight topics belongs to four different modes of discourse. The observed significant result shows the difference between individual prompts and it is necessary to calculate the total mean for each discourse type by aggregating the scores for paired topics. The results of these calculations are given in Table 11, which apparently shows a clear order.

Table 11: Aggregated mean scores for topics paired for the discourse type

Variable	Mean	Std Dev	Minimum	Maximum	N
ARG	5.78	2.08	2.00	10.00	87
EXP	6.74	2.28	2.00	10.00	87
NAR	7.09	2.10	2.00	10.00	87
DES	7.13	2.02	2.00	10.00	87

In the next step of the statistical analysis, a repeated measures ANOVA was run to ensure that the apparent rank order of difference

among the means of these four discourse types is significant. Table 12 shows the result:

Table 12: The result of repeated measures ANOVA for different discourse types

Source of Variation	SS	DF	MS	F	Sig of F
WITHIN+RESIDUAL	1096.91	258	4.25		
DISCOURSE TYPE	102.59	3	34.20	8.04	.000***

As the result of this analysis, it might be claimed that there is a significant difference between the students' perception of these four discourse types ($p < 0.000$). Still, it is necessary to find the location of the difference by running the Tukey's HSD multiple range test. Table 13 presents the location of significant differences with asterisks.

Table 13: The Tukey's HSD test for the total means of paired topics

Mean	Mode	nar	des	exp	arg
7.09	nar				
7.13	des				
6.74	exp				
5.78	arg	*	*	*	

As it is obviously projected in Table 13, argument is different from all other discourse types. Although there is considerable difference between means for explanation and narration (0.35) and also between means for explanation and description (0.39), these differences, while more relaxed post hoc tests may find them significant, are not, in fact, big enough to be considered significant by a much more conservative test like Tukey's HSD which requires a mean difference of (0.61).

Considering the fact that the repeated measures analysis of variance for both individual prompts and discourse-type-based paired prompts shows a significant difference, it can be concluded that the fourth null hypothesis is rejected and hence the alternative hypothesis is accepted:

Hypothesis #4: "THERE IS a significant difference between the learners' choice of essay topics which will elicit texts of different discourse natures."

It should be pointed out that the difference implied in this hypothesis needs to be restricted to argument and the rest of discourse modes, as located by the Tukey's HSD test. However, it was quite likely to find a wider range of difference among modes, if more lenient post hoc multiple-range tests were applied.

Discussion

Argumentative essay prompts once more proved to be the most difficult discourse type and therefore received the lowest ratings. Next to Argument and second in order of difficulty is Explanation. And quite interestingly, Description and Narrative seems to be the most favorite types of compositions.

Narrative #2, for the reasons discussed earlier is exceptional. This is a warning to the makers and users of tests not to rely on native-speakers and their intuitions in the process of topic selection. Hamp-Lyons and Mathias (1994) have already pointed out that experts' judgments about topic difficulty are not reliable. However it seems that the learner's judgments are more accurate. As the issue of familiarity is culture-bound, it is still not clear how other L2 learners with different cultural backgrounds react to this sort of discrepancy.

CONCLUSIONS

According to the results of the present study, which has been an attempt to fill in the existing gap in the field of second/foreign language writing with regard to the impact of mode of discourse on essay topic difficulty, the following conclusions may be drawn:

1. A significant difference among essays elicited in different modes of discourse has been observed. Argument is clearly more difficult to produce. Narrative, on the other hand, seems to be the easiest one. Description stands in between.
2. Exposition does not follow any predictable pattern, This could be due to its more complicated nature which may require a combination of other discourse modes to write an expository

- essay. Further research is recommended to clear the situation by breaking exposition into smaller sub-modes.
3. The difference among the four discourse types is not simply a matter of training and practising in writing across these modes. They are linguistically and cognitively different.
 4. Students are consciously or sub-consciously aware of such differences and prefer easier modes over difficult ones.
 5. Mode of discourse is just one of the several variables affecting the difficulty level of topics. There are other variables, some of which even may have an overriding effect. Further research is required to investigate the weight of these affecting variables.

Received 11 January 2003

Accepted 30 July 2003

REFERENCES

- Applebee, A. (1982). Writing and learning in school setting. In Nystrand M. (Ed.) *What writers know: The language process and structure of written discourse*. New York: Academic Press.
- Braddock, R., Lloyd-Jones, R., and Schoer, L. (1963). *Research in written composition*. Urbana, IL: National Council of Teachers of English.
- Bridgeman, B. and Carlson, S. (1983). A survey of academic writing tasks required of graduate and undergraduate foreign students. TOEFL Research Report No.15. Princeton, New Jersey: Educational Testing Service.
- Britton, J., Burgess, T., Martin, N., McLeod, A., & Rosen, H. (1975). *The development of writing abilities (11-18)*. London: Macmillan.
- Brossell, G. & Hoetker Ash, B. (1984). An experiment with the wording of essay topics. *College Composition and Communication*, 35, 3, 423-25.
- Carlson, S., Bridgeman, B., Camp, R., & Waanders, J. (1985). *Relationship of admission test scores to writing performance of native and nonnative speakers of English* (TOEFL Research Report 19). Princeton, NJ: Educational Testing Service.

- Chiste, K. & O'Shea, J. (1988). Patterns of question selection and writing performance of ESL students. *TESOL Quarterly*, 22, 4, pp. (6)281-(6)284.
- Crowhurst, M., & Piche, G.L. (1979). Audience and mode of discourse effects on syntactic complexity in writing at two grade levels. *Research in the Teaching of English*, 13, 101-110.
- Davies, A. (1965) *Proficiency in English as a second language*. PhD Thesis. University of Birmingham.
- Dudley-Evans, T. (1988). A consideration of the meaning of "discuss" in examination questions. In P. Robinson (Ed.), *Academic writing: Process and product: ELT Documents No. 129*, (pp.47-52). Hong Kong: Modern English Publications and The British Council.
- Farhady, H., Jafarpoor, A. and Birjandi, P. (1995) *Testing Language skills: from theory to practice*. Tehran: SAMT Publication.
- Freedman, S. W. and Calfee, R. C. (1983) Holistic assessment of writing: Experimental design and cognitive theory. In Mosenthal, P., Tamor, L., and Walmsy, S. (eds.), *Research in writing: Principles and methods*. Longman, pp. 75-98.
- Hamp-Lyons, L. (1986). *Assessing second language writing in academic settings*. Unpublished doctoral dissertation, University of Edinburgh, Scotland.
- Hamp-Lyons, L. (1988). The product before: Task-related influences on the writer. In P. Robinson (Ed.), *Academic writing: Process and product: ELT Documents No. 129*, pp.35-46. London: Macmillan/British Council.
- Hamp-Lyons, L. & Kroll, B. (1997) TOEFL 2000- writing: composition, community, and assessment (TOEFL Monograph Series). Princeton, NJ: Educational Testing service.
- Hamp-Lyons, L. & Mathias, S.P. (1994) Examining expert judgments of task difficulty on essay tests. *Journal of Second Language Writing*, 3, 1, pp. 49-68.
- Hamp-Lyons, L. (1990) Second language writing: Assessment issues. in Kroll, B. (ed.) *Second language writing: research insights for the classroom*. New York: Cambridge University Press, pp.69-87.

- Hartog, P. et al. (1941) *The marking of English essays*. London: Macmillan.
- Hirokawa, K., & Swales, J. M. (1986). The effects of modifying the formality level of ESL composition questions. *TESOL Quarterly*, 20, 343-345.
- Hirsch, E. and Harrington, D. P. (1981). Measuring the communicative effectiveness of prose. In Fredrikson, C. and Dominic, J. (eds.) *Writing: The nature, development and teaching of written communication. Vol.2: Writing: Process, development and communication*. Hillsdale, N. J.: Erlbaum. pp. 189-207.
- Keller-Kohen, D. & Wolfe, A. (1987). *Extending writing in the college of literature, science and the arts: Report on a faculty survey*. Ann Arbor: University of Michigan English Composition Board.
- Kincaid, G. L. (1953). *Some factors affecting variations in the quality of students' writing*. Unpublished doctoral dissertation, Michigan State University, East Lansing, MI.
- Nemati, M. (1996). The effects of discourse on EFL/ESL writing performance. In M. Cortazzi, S. Rafik Galea, and B. Hall (Eds.), *Aspects of language teaching, learning and research methodology in the context of education*, Leicester: University of Leicester, pp. 162-167.
- Park, Y. M. (1986). *The influence of the task upon writing performance*. Unpublished doctoral dissertation. University of Illinois, Urbana-Champaign.
- Perron, J. D. (1977) The impact of mode on written syntactic complexity, Parts 1-3. *Studies in Language Education: Reports* 24, 25, 27. Athens: Department of Language Education, University of Georgia.
- Poetker, J. S. (1977) Practical suggestions for improving and using essay questions. *High School Journal*, 61, pp. 7-15.
- Pollitt, A. Hutchinson, C. Entwistle, N. and DeLuca, C. (1985) What makes exam questions difficult? An analysis of "O" grade questions and answers. *Research Reports for Teachers*, No.2. Edinburgh: Scottish Academic Press.

- Quellmalz, E., Cappell, F. and Chou, C. (1982). Effects of discourse and response mode on the measurement of writing competence. *Journal of Educational Measurement*, 19, pp. 241-258.
- Raimes, A. (1990). The TOEFL test of written English: Causes for concern. *TESOL Quarterly*, 24, 3, pp. 427-441.
- Reed, W. M., Burton, J.K. & Kelly, P.P. (1985) The effects of writing ability and mode of discourse on cognitive capacity engagement. *Research in Teaching of English*, 19, 3, 283-297.
- Reid, J. M. (1990). Responding to different topic types: A quantitative analysis from a contrastive rhetoric perspective. In Kroll, B. (ed) *Second language writing: Research insights for the classroom*. Cambridge: Cambridge University Press, pp. 191-210.
- Rosen, H. (1969) *An investigation of the effects of differentiated writing assignments on the performance in English composition of a selected group of 15/16 year old pupils*. Ph.D. Thesis, University of London.
- Ruth, L., & Murphy, S. (1988). *Designing writing tasks for the assessment of writing*. Norwood, N.J.: Ablex.
- Scott, V. M. (1996). *Rethinking foreign language writing*. Boston, Mass: Heinle & Heinle Publishers.
- Spaan, M. (1989). *Essay tests: What's in a prompt?* Paper presented at the 23rd Annual TESOL Convention, San Antonio, Texas, March.

APPENDIX 1

Descriptive topics

#1 "Imagine that you are in the 25th century. Describe the people and their way of life, buildings, vehicles, etc."

#2 "Describe your hometown or any interesting city so that readers can have a clear image of the city in their minds."

Argumentative topics

#1 "Argue the advantages and disadvantages of commercial advertising in the media. Take a clear position."

#2 "At the present time, the human race spends huge sums of money to explore space. Do you find it a waste or do you view space projects as essential to expanding our knowledge. Discuss the issue."

Narrative topics

#1 "Write the story of the most interesting film you have ever seen. Narrate the events as they happened in the film. Your writing should look like a story."

#2 "Write the story of someone who rose from an unknown background and humble beginning to become successful and famous. Narrate the main events of his/her life story in the order in which they happened."

Explanatory topics

#1 "How can we make teaching and learning in Iranian universities more effective?"

#2 "Considering the situation in Iran, explain how a foreign language should be learned?"

APPENDIX 2

Exploring the space or any other scientific undertaking in itself is neither good nor bad. It depends on the ends it h in view and the circumstances which e its immediate enactment or its postponement f times more appropriate, or even leaving i a mere theory or suggestion untested b practice. Of the last, we can t of many theories that putting them i practice can never bring any advantage t man or, in some circumstances, only a to his misery. Programmes of space e are of the second kind which t potentially entail no harm to man a even could be of much help i solving the riddles of life and o of existence, yet their benefit depends o the current circumstances and the conditions in which man lives. In cases like these, we should act according to the maxim "first things first".

If you come to my city from a long distance in your car, you can see the green gardens which surround it since it is located in a higher place than its surrounding. It is not an ordinary city b originally a big oasis in the l friendly environment of the desert. At t border of the city where the s spreads, you can see the small s and bushes which are called by t locals the fiery tongues because their gl narrow leaves reflect the sunshine and a as narrow streaks of light or f during the day. The earth under t bushes is covered by mat o thick green grass which to the c to the city appears a beautiful s that distinguishes the yellow brown sand f the green floor of these bushes. H passed these shrubs which are here and there, you can see young men with their double-barrelled hunting guns. Then you come to a wide extent of acres of field planted with wheat, barley, oats, as these are the city's major agricultural products.

"Ghost" is one of the interesting films I saw two years ago. The main character of the film i a young man who is honest, h and hardworking. Along with his friend h works in one of big companies a an accountant or something like this. H

girlfriend is a slim, innocent-looking young woman with big eyes and short hair. She is an artist. One day the hero, I mean the main character, discovers that he has millions of dollars in his account and shares his secret with his colleague. The friend turns out to be a thief and at night, while the hero is walking in a dark street with his girlfriend, he sends an ugly-looking villain to kill him so that he can't take the whole money for himself. The hero gets killed after a futile fighting with the villain.

Learning a foreign language, like any other activity, arises from and is motivated by the need which man feels for a foreign language. Once the motivation is strong enough, one would look for those procedures and techniques that will satisfy his need. There is no doubt that language learning should include all the four skills of learning, yet the motivation for the learning is not the same for all those who learn the language. Some people would find speaking is of their major importance, while others' preference could be of any of the other three skills. Therefore, learning procedures should not be kept stiff and inflexible in all situations. The procedures also should vary according to age group, sex, and social and cultural background. Depending on the age, the method and appropriate techniques will change. Naturally, the inclination with an adult will be most for reading rather than speaking which is more favored by younger ages.