Phonological Metathesis in Persian: Synchronic, Diachronic, and the Optimality Theory

Mohammad Reza Ahmadkhani*
Associate Professor, Payame Noor University, Varamin, I.R. Iran
(Received: 22 July. 2008, Accepted: 10 Oct. 2009)

Abstract
Metathesis as a phonological phenomenon in order to reverse segments has been focus of many studies. Metathesis according to the range of application can be sporadic or regular. It links synchronic studies of language to diachronic ones and thereby has some consequences for phonological theory. The central goal of this paper is to elucidate the conditions under which metathesis occurs synchronically in modern Persian and diachronically in Pahlavi (Middle Iranian Persian). In this paper, first metathesis is viewed from different aspects especially a historical one. And particular attention is paid to its causes and contexts. Then, metathesis is studied within Optimality Theory, using constraints such as Syllable Contact Law (SCL), Sonority Sequence Principle (SSP),… Also it was found that segments don't behave identically in the metathesis process. Phoneme /r/ is the most frequent and /z/ is the least one found in Persian.

Key Words: Metathesis, Persian, Constraint, Optimality Theory.

* Tel: 0292-21157111, Fax: 0292-2114282, E-mail:ahmadkhani@pnu.ac.ir
1. Introduction

One of the long standing issues in linguistic theory is metathesis. Metathesis can be defined as the process where by in certain languages, under certain conditions sounds appear to switch positions with one another. Thus in a string of sounds where we would expect the linear ordering of two sounds to be …xy…, we find instead…yx…..

The explanation to the metathesis has been to say that this process is indeed less natural phonetically than other processes, and has a relatively greater phonological motivation. First suggested by neogrammarian movement (Osthoff & Brugman 1878: xvi), this view was clearly stated by Brugmann (1902:246): "Metathesis arises when the order of sounds and the syllable boundary make inconvenience; it causes a group of sounds to be placed where it is easier for the speaker". This phonetic optimization approach has prevailed in descriptive and typological studies. According to Grammont (1950:239), for instance, metathesis yields a better syllable structure safeguards unity and harmony of languages sound system in replacing unusual groups by common groups which are or have become unpronounceable in substituting simple types for them and avoids useless articulatory effects. Últan's (1985:395) survey also concludes that the superficial cause of most metatheses is conversion of a phonologically inadmissible or disfavored sequence into an acceptable one. Finally, Hock (1985:532-33) contends that metathesis can become regular only when it serves a specific structural purpose usually that of converting phonologically or perceptually marked structures into more acceptable ones.

2. The Nature of metathesis

Metathesis has resisted a unified, explanatory treatment in nonlinear phonology despite advancements in the formalism used to account for many other processes, such as assimilation (Clements, 1985) and dissimilation (Odden 1987). Unlike these phenomena, there is no unique formalism for characterizing metathesis as a primitive rule type. Instead metathesis effects have been derived by a variety of different means including: successive application of rules of deletion and insertion...
Phonological Metathesis in Persian ...


Sound changes mainly are caused by listener-based reinterpretation. This in turn may arise in several ways, for example, the actual phonetic string may simply misperceive the utterance due to speaker variation on a continuum from hyperarticulated listener oriented clear speech to reduced hypoarticulated casual speech. To be exact acoustic/auditory similarity between sounds can have the effect of diminishing the degree of distinctiveness of the sound thereby making them and their order less easily identifiable (Hume 1998). So ease of production and perception are motivations for phonological processes.

Phonetic cues and contexts are crucial for identification of speech sound. For example when intervocalic stop and fricative are involved, the stop consistently surfaces before vowel (Steriade, 2001). This claim accounts for metathesis in Udi (see Salvia 1973; Makashay, 2001).

1. Udi
/tad – esun/ [tastun] “to give”
/bafd –sa/ [bafsta] “falling into”

With respect to word position, it has been claimed that the left edge of the root or word is special for word recognition (Beckman 1998; Cutler et al., 1985; Marslen-Wilson, 1989; Marslen-Wilson and Switserland, 1989), since lexical access is generally achieved on the basis of the initial part of the word. Hence, salient information tends to occur at the beginning of words, and this location tends to resist phonological processes. Observed patterns of metathesis are consistent with this over all patterns.

Perceptual optimization may also play a key role in shaping patterns of metathesis. As discussed in Hume (1998, 2000), metathesis frequently occurs in contexts of low salience and serves to enhance the contrast of the sounds in that context. Thus the form metathesis is superior to the expected unmetathesised form in
terms of the overall perceptual salience of the segments involved. This can be
achieved by shifting a consonant from an environment in which the clues to its
identification are more robust, or by reordering segments in order to improve the
overall salience of a neighboring sound or sounds.

Ultan (1978; 383-90) mentions some conditions for occurrence of metathesis:
a) reduction as in apocopy of vowels in Persian: suxr (Pahlavi) > surx (Persian)
b) open syllable canon, as in French, vervecem > berbis (young)
(also see Martinet, 1955:349-356)
c) phonetic constraint such as eliminating initial or final clusters with liquids as
setl (Old English) > seld (English)
d) attraction, such as drit, brid (middle English) > dirt, bird (Modern English),
in which although Cr cluster is admissible, post vocalic dental exerted a strong
attraction on the r.

Alexander (1985: 34-5) formalize this attraction as following:
2. [- coronal] r v [+ coronal ] → 1,3,2,4

Thereby he suggests a hierarchy of consonants (3) to be drawn up, with the left
most consonants being the ones strongest in the attraction of r. The notation “ >”
when used in (3) below, means ‘is stronger than’:
3. d > l > θ > k > f > h
t s g w
n p
b
m

Hock (1985:534) suggests that regular metathesis serves to eliminate clusters
when it fails to conform to the preferred syllable structure according to which
sonority increases in the onset and decreases in the coda. According to Selkirk
(1984) the Sonority Sequence Principle in (4) specifies the order in which
consonants arrange themselves from the syllable nucleus to its edge:
The proneness of different phonetic classes to metathesis tends to stand in direct correlation with a hierarchy of resonance. By increasing the sonorant quality of a sound, its tendency to undergo the metathesis also increases. The sole exception involves an original sequence of voiceless stop + sibilant, metathesis of which was found to be fairly common and widespread. In Armenia, the evidence points to a chronological hierarchy in the introduction of metathesis of original clusters of the type consonant + semi vowel. The first to metathesis were clusters containing semivowels, followed by the liquids, nasals, spirants, stops, and possibly the affricates, in that order (Ultan 1978:395). It is governed phonotactically, according to Grammont (1923) that less sonorant consonants (those with smaller aperture) are always positioned closer to syllable boundary and more sonorous consonants closer to syllabus nucleus.

The disproportionaly high (and widespread) frequency of occurrence of liquids in metathesis is proverbial. A partial list of language for which this is true includes:

- Liquids: Breton, Cornish, Eastern Eskimo, Middle and Old French, Gaelic, Ancient Greek, Old Icelandic, Indo-European, Indonesia, Old Irish, Latin, Mandaic Aramic, Persian, South Slavic, Old Spanish, Tagalog, and Zoque.
- R only: Anglo-Norman, Armenian, Avestic (and Zend), Bagneres -de- Luchon, Middle English, Italian (Supraselva), Kamhmu, Mauritian Creole, Vedic Sanskrit, Sardinian, and Toba.
- L only: Amuzgo and Yokuts (Ultan, 1978:392)

Hume (1997:153-56) explains some motivations for metathesis in Leti. Metathesis before a consonant cluster serves to avoid a complex syllable margin. Thus, the surface syllabification of a form such as /ulit prail/ is claimed to be ul.tip.ra.i with the first consonant of the second morpheme syllabifying as coda of the preceding syllable. Were metathesis not to apply, we would expect the initial consonant of the second morpheme to syllabify as part of a complex syllable margin, e.g. *u.lit.pra.i or *u.litp.ra.i., more examples:

- Leti
  - /ukar+ppalu/ ukrappalu ‘index finger’
  - /danat+kviali/ dantkviali ‘millipede’
Support for this claim comes from this fact that before a morpheme beginning with a simple onset, as in the followings, a consonant final morpheme does not undergoes metathesis:

Leti

/urun/ urun moa
/lout/ lout de

Another function of metathesis in Leti is the avoidance of onsetless syllables. In /ukar +muani/ > ukramwan metathesis occurs before a morpheme initial /CVV/ sequence, where the final vowel and consonant of the initial morpheme reverse positions. In this case metathesis is motivated by the need to avoid vowel hiatus: that is, an onsetless syllable:

In this section it was argued that metathesis arises to have a better syllable structure to produce and perceive speech sounds. Also for accounting metathesis the nature of sound and the context of occurrence are important.

3. Types of metathesis

Metathesis is categorized into four types: perceptual metathesis, compensatory metathesis, articulatory metathesis, auditory metathesis.

3.1. Perceptual metathesis

Perceptual metathesis in which certain features in effect jump over adjacent segments is closely linked to the segmentation problem. Perceptual metathesis originates when features, extending across a CV or VC domain or perceived as extending across such a domain, are perceived as originating in a position which is not the surface position. As emphasized by Ohala (1993) in his discussion of dissimilation, certain perceptual features are typically realized over relatively short time duration, whereas others are typically realized over relatively long duration. For example, irrespective of its phonological association with a consonant, vowel or glide, pharyngealization is typically phonetically realized over a minimal CV or VC domain. In case of perceptual metathesis, a segment (or feature) with elongated
phonetic cues shifts its linear position in a phonological string. This fact partly reflects the perceptual difficulty of localizing the origin of phonetic cue with long-distance effects (Blevins and Garrett 1998:510).

The result of perceptual metathesis is a mistake from the point of view of the previous linguistic system; a segment (or feature) is reinterpreted as originating in a new position within the elongated span. This will involve the transposition of adjacent elements in some cases, and in other cases metathesis will be non-local. Examples from Cayuga (Foster 1982) are in (7):

(7) Cayuga
   a. /kahwista?eks/       [khawisd?aes]
   b. /akekaha?/           [agekhaa?]

A comparable sound change has occurred in Rendille (a Cushitic language in Kenya) and is still manifested in synchronic alternation involving underlying obstruent or and nasal- r sequence (Heine 1926, Oomen 1981, Sim 1981).

(8) Rendille
   agar-te  (2sg=3sa.FEM)                             arg-e (1SG=3SG.MASC)
   ugar   (singular)                                         urg-o (plural)

Long distance liquid metathesis has occurred as a sound change in South Italian dialects of Greek (Rohlfs 1924). In this dialect, prevocalic or in a non-initial syllable has been transposed into the initial syllable in certain circumstances. This occurred whenever (i) the liquid was positioned after an obstruent and either (ii) a the initial syllable had a prevocalic non-coronal obstruent or (ii) b the liquid was r and the initial syllable had a prevocalic l. If these conditions were satisfied, the liquid moved into prevocalic position in the initial syllable as in Greek (9):

(9) Classic Greek                   South Italian Greek
    kopros                                               kropo     “dung”
    gambros                                              grambo   “son-in-law”

However following examples show that (10a) the intervocalic and preconsonantal liquids are unaffected (10b) liquids are transposed only into initial syllables with prevocalic obstruents (10c) that metathesis never yields clusters
consisting of a coronal obstruent plus / or consisting of a coronal fricative or affricate plus r.

(10) Classic Greek South Italian Greek
a. kalos kalo “attractive”
b. metron metro “measure”
c. seklion sekli “beet greens”

Perceptual metathesis involving labialization and palatalization is also well attested. Comparable long-distance cases are found among the Ethiopian Semitic labialization and palatalization processes described by Hetzron (1971, 1977:45-49).

Perceptual metathesis involve pharyngeals. A local pharyngeal metathesis is said to exist in Rendille (11), where the pharyngeal fricative switches with an adjacent consonant when preceded by the low vowel (Hume 1997:294):

Rendille

<table>
<thead>
<tr>
<th>non-prevocalic</th>
<th>prevocalic</th>
</tr>
</thead>
<tbody>
<tr>
<td>aham (sg)</td>
<td>amh-a (pl)</td>
</tr>
<tr>
<td>Bahab (sg)</td>
<td>babh-o (pl)</td>
</tr>
</tbody>
</table>

Glottalization metathesis is found in some languages. The following examples (12) from Interior Salish language, Shurwap, show various forms associated with a single suffix containing a glottalized sonorant. In this language nonsyllabic glottalized sonorants do not surface in postconsonantal position. If an underlying glottalized sonorant is post consonantal and to the right of the main accent, its glottalization shifts leftward unto an immediately post tonic sonorant, if there is one (Kuipers 1974).

(12) Root and suffix Derived forms
-él'qs “clothing” t-kwltk-él'qs “underwear”
qwey- "black" qwey'-lqs "priest"

The table (1) shows consonant types attested in perceptual CV metathesis (Blevins and Garrett, 1998: 513):
Table (1)

<table>
<thead>
<tr>
<th>segment</th>
<th>feature</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>liquids</td>
<td>laterality</td>
<td>Slavic-Bagneres-De-French</td>
</tr>
<tr>
<td>laterals</td>
<td>laterality</td>
<td>Latin</td>
</tr>
<tr>
<td>rhotics</td>
<td>rhoticity</td>
<td>Le Harve French</td>
</tr>
<tr>
<td>laryngeals</td>
<td>aspiration</td>
<td>Cherokee</td>
</tr>
<tr>
<td>h</td>
<td>glottalization</td>
<td>Zoque,Mohawk</td>
</tr>
<tr>
<td>?</td>
<td>pharyngealization</td>
<td>Proto-Indo-European, Arabic dialect</td>
</tr>
<tr>
<td>pharyngeals</td>
<td></td>
<td>Koshin,Birom,Leti</td>
</tr>
<tr>
<td>glides/vowels</td>
<td>palatalization</td>
<td>Greek</td>
</tr>
<tr>
<td>j/i</td>
<td>labialization</td>
<td>Aghem,Noni</td>
</tr>
</tbody>
</table>

3.2. Compensatory metathesis

Compensatory metathesis, as the name suggest, is prosodically conditioned. In this type of metathesis a vowel at the edge of the phonological domain undergoes phonetic weakening in quality and duration, with compensation for this weakening by anticipatory or perseverative coarticulation of the original vowel quality in stressed position. The sequence of sound changes occurring at each end of the relevant phonological domain is shown in (13):

(13) right edge: $V_1CV_2>...V_1V_2CV_2>...V_1V_2C$
    left edge: $V_1CV_2>...V_1CV_1V_2>...CV_1V_2>...$
where a vowel at the edge of the phonological domain undergoes phonetic weakening in quality and duration, with compensation or preservative coarticulation of the original peripheral vowel quality in non peripheral stressed position. Followings are examples from Rotuman and Nako (Blevins and Garrett 1998:527):

(14) Rotuman:
- futi  fyt
- tiko  tiok

3.3. Coarticulatory metathesis

Coarticulatory metathesis is a type of metathesis with articulatory origin. Extreme coarticulation is possible in a sequence of stops, each of which involves closures of a distinct articulator. When $C_1C_2$ gestural overlap results in nearly simultaneous closure, with $C_2$ realized after $C_1$, a $C_2C_1$ cluster may be perceived. One of these types is labial velar stop sequence. In Bisayan languages, Cebuano and Aklanon,(15) we see this kind of metathesis (Zorc 1964:97):

(15) Cebuano   Aklanon
- libgus   ligbus      "mushroom"
- palibga  paligba    "confuse him"

In this type there is coronal-non coronal stop sequences, examples from Cebuano Bisayan. Data is cited below (16) from Blust (1979:110).

(16) metathesis   no metathesis
- nm inum:imn-a “drink”     mn daman:damn-un “talk”

3.4. Auditory metathesis

Auditory stream decoupling leads to metathesis involving sibilant + stop and stop + sibilant metathesis. Some Examples from West Saxon dialect of Old English are in (17) (Weyhe 1908; Campbell, 1959:177-78; Luick, 1921:913-14; Jordan, 1974:168-70). In this dialect $sk$ clusters regularly inverted their linear order and
became ks clusters.

(17) Old English                             Late West Saxon
     frosk                                   froks
     husk                                    huks
     aske                                    akse
     fiskas                                  fiksas

According to Grammont (1923:73) a ks > sk change has occurred word-finally in colloquial French.

(18) French
     Standard                               Colloquial
     fiks                                    fisk
     lyks                                    lysk

Blevins and Garrett (1998) attribute the reverse order of metathesis in Old English and French to longer sibilants in final positions in French which causes greater confusion effect on segmental order. This is based on observation that French accentual -phrase-final syllables are significantly longer than non-accentual-phrase-final syllables.

In this section different types of metathesis were examined: a) perceptual metathesis in which features of one segment is transmitted to other segments, b) compensatory metathesis in which omission of a sound and compensating it prosodically leads to metathesis c) coarticulatory metathesis in which easy articulation and overlapping of segments causes metathesis and d) auditory metathesis in which auditory stream decoupling leads to metathesis.

4. Metathesis in Persian

4.1 Synchronic metathesis

In Persian, metathesis is found in child language, speech error, and colloquial speech of illiterate people recorded by the author.

As the following examples (19) from Modern Persian show, liquid and sibilant
sounds metathesis with stops (k,t,b,m,q) and fricative (j,f). In (19a) the motivation for metathesis is to observe Sonority Sequence Principle (SSP) and in (19) the motivation is to observe Syllable Contact Law (Hooper, 1972), according to which in heterosyllabic cluster of $C_1 C_2$ there is the tendency of $C_1$ to be more sonorant than $C_2$.

(19) Persian

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>luks</td>
<td>lusk</td>
<td>“luxurious”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>madrese</td>
<td>mardese</td>
<td>“school”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The skeletal model of metathesis in *lusk* is shown in (20):
4.2. Diachronic metathesis

Diachronic metathesis includes examples (21) from Pahlavi (Boyce, 1977; Nyberg 1931) in which consonant + liquid undergoes metathesis to be liquid + consonant to follow Sonority Sequence Principle. In (21a) apocope of the final vowel in the final syllable provides the context for metathesis.

(21) Pahlavi Persian

<table>
<thead>
<tr>
<th>Pahlavi</th>
<th>Persian</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. suxra</td>
<td>surx</td>
</tr>
<tr>
<td>žafra</td>
<td>žarf</td>
</tr>
<tr>
<td>wafra</td>
<td>barf</td>
</tr>
<tr>
<td>asru</td>
<td>ars</td>
</tr>
<tr>
<td>čaxra</td>
<td>čarx</td>
</tr>
<tr>
<td>namra</td>
<td>narm</td>
</tr>
<tr>
<td>b. taxl</td>
<td>talx</td>
</tr>
<tr>
<td>hagriz</td>
<td>hargez</td>
</tr>
<tr>
<td>wrata-</td>
<td>rvata-</td>
</tr>
<tr>
<td>wrinati-</td>
<td>rvinati</td>
</tr>
<tr>
<td>haqle</td>
<td>halqe</td>
</tr>
<tr>
<td>laqtan</td>
<td>qaltan</td>
</tr>
</tbody>
</table>

In the followings, (22) consonant cluster Cr of the first syllable of Pahlavi words eliminated by metathesis of liquid r with vowels to avoid complex onset (Prince and Smolensky, 1993). In the examples, r is attracted by d, h and j in (a)-(c), and by z and y in (d) and (e).
Here, changing CCV pattern in Pahlavi to CV in initial syllable of words in Persian as an attempt to avoid complexity (CC) in the syllable onset (Prince and Smolensky, 1993) is the motivation for the metathesis. On the basis of the data in (19),(21) and (22), figure (1) shows the frequency of metathesised phonemes in Persian words. The high frequency of \( r \) is related to the articulator, tongue, which is the most moveable and dynamic speech articulator. Also, according to the figure, \( r \) is more frequent than \( l \) since it is produced by the tip of the tongue which is more moveable than the blade of the tongue which is involved in the articulation of \( l \).
4.3. The Optimality Theory

Ever since Optimality Theory (Prince and Smolensky 1993) came into circulation, it has been perceived by a considerable number of generative phonologists worldwide, as perhaps current as the most promising framework of analysis. The theoretical framework of this thesis is that of OT, mainly because of OT’s ability to capture universal properties and select outputs in a manner which reflect crosslinguistic tendencies as well as predict typological variations. OT’s way of capturing generalizations is through conflicting constraints, and an OT grammar of a language is expressed in the ranked order of violable constraints.

Optimality theory is based on the following assumptions: Languages, or better, the words and phrases in individual languages are shaped by the interaction of universal constraints on output forms. These universal constraints stand in conflict and they are ranked with regard to each other. The notion of constraint conflict means that it is often only possible to satisfy one constraint on the coast of violating another. The notion of constraint ranking is more important than satisfying others.

In Optimality Theory, grammar consists of the following components (Prince & Smolensky 1993): Con (constraint), a set of violable universal constraints, ranked on a language-particular basis, against which the well-formedness of output candidates is evaluated; a function Gen, which associates an input form with a potentially infinite set of output candidates and orders them according to how well they satisfy the constraint system of the language in question. The actually occurring output form is that candidate which best satisfies the constraint system.

According to the optimality theory and constraints such as Linearity (no metathesis) (McCarthy and Prince, 1993, 1999), Complex Onset Avoidance (*Complex) (Prince and Smolensky, 1993), Syllable Contact Law (SCL) and Sonority Sequence Principle (SSP) the output of *luks*, *kebrit*, and *draya* are represented respectively in tables (2),(3),and (4). The stars in the cells show that the word violates the correspondent constraint and is not optimal. Although in the tables words violate both constraints, but violation of the higher (left) constraint is fatal (!) and crashes the word.
In table (2) option (b) is optimal since it violates the Linearity constraint which is lower in the rank than SSP constraint.

Table (2)

<table>
<thead>
<tr>
<th></th>
<th>Input: /luks/</th>
<th>SSP</th>
<th>Linearity</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>luks</td>
<td>!*</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>lusk</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

In table (3) option (b) is optimal since it violates the Linearity constraint which is lower in the rank than SCL constraint.

Table (3)

<table>
<thead>
<tr>
<th></th>
<th>Input:/kebrit/</th>
<th>SCL</th>
<th>Linearity</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>kebrit</td>
<td>!*</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>kerbit</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

In table (4) option (b) is optimal since it violates the Linearity constraint which is lower in rank than *Complex constraint.

Table (4)

<table>
<thead>
<tr>
<th></th>
<th>Input:/draya/</th>
<th>*Complex</th>
<th>Linearity</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>draya</td>
<td>!*</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>darya</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

5. Conclusion

The aim of this paper was giving an account of accounting metathesis in Persian. In the synchronic study data showed that in one syllable words, the motivation for metathesis is observing Sonority Sequence Principle with
phonological tendency of s,l,r,z to be in the center of the syllable and less sonorant sounds in the coda of syllable and thereby observed Sonority Sequence Principle. In the two-syllable words for improving phonetic cues and perception, there is tendency of observing the Syllable Contact Law. In diachronic metathesis besides SSP, the onset Cr cluster is eliminated by metathesis of r with the following vowel to have simple onset which shows that explanation of the same phenomenon could be different in synchronic and diachronic studies. Regarding attraction of r by other segments, it was shown that attraction of coronals is stronger than semivowels and attraction of semivowels is stronger than glottal and palatals.

These generalizations show that metathesis is not an end but a means for ease of perception and production. Of course there are other constraints such as social labels which control spreading of metathesis to all potential ones to avoid pragmatically unacceptable structures. Furthermore, the tendencies contribute to claim that at least one way of accounting synchronic phonological structure is using typological generalizations resulting from the pathways of historical changes. In other words if historical change operates in such a way to favor or disfavor certain situations, its results are what we will find and such generalizations could be guides to the structure of the language faculty.

References
Besnier, Niko. (1987)."An autosegmental approach to metathesis in Rotuman". Lingua 73.


Iranian English Teachers’ self-efficacy: Do Academic Degree and Experience make a difference?

Ramin Akbari*
Assistant professor of Tarbiat Modares University, I.R. Iran

Shahab Moradkhani**
MA student of Teaching English as a Foreign Language (TEFL) at Tarbiat Modares University, I.R. Iran

(Received: 22 July. 2008, Accepted: 10 Oct. 2009)

Abstract
Teacher efficacy refers to teachers’ beliefs in their ability to enhance student achievement and bring about positive learning outcomes. The present study was intended to investigate possible relationships between experience/academic degree and teacher efficacy among EFL teachers. Four hundred and forty-seven teachers who participated in this study filled in a survey which included some demographic information as well as Teacher Self-Efficacy Scale (TSES). The results of data analysis showed that experienced teachers (with more than three years of teaching experience) had a significantly higher level of global efficacy, efficacy for student engagement, efficacy for classroom management, and efficacy for instructional strategies compared to their novice counterparts. In contrast, teachers who had English-related academic degrees did not enjoy significantly higher levels of efficacy except in the subcomponent of student engagement. The findings are discussed in the light of different sources of efficacy to which novice/experienced teachers resort and the nature of English-related university programs.

Key Words: Teacher Efficacy, Teacher Experience, Teachers’ Academic Degree, Novice Teachers, Experienced Teachers.

* Tel: 021-82883664, Fax: 021-82883669, E-mail: akbari_ram@yahoo.com
** Tel: 021-66908252, Fax: 021-82883669, E-mail: sh_moradkhani2000@yahoo.com
1. Introduction

Before the 1990s, despite the fact that teachers had always constituted one of the main rings of the educational chain, they had been neglected to a large extent in the research agenda. However, in recent years, with the postulation of postmethod pedagogy which empowers language teachers “to theorize from their practice and practice what they theorize” (Kumaravadivelu, 2001, p. 541) and critical pedagogy which considers teachers as “transformative intellectuals” (Pennycook, 1989, p. 613), more attention has been paid to them. Consequently, some researchers have investigated different characteristics of language teachers such as their pedagogical knowledge base (e.g., Watzke, 2007), professional development (e.g., Ross & Bruce, 2007), and identity (e.g., Tsui, 2007) which affect teachers’ classroom practices and subsequently students’ achievement.

One of the features that has absorbed a good deal of attention recently is teachers’ sense of self-efficacy which is a crucial parameter in determining teachers’ opinion about their job, their classroom activities, and their influence on students’ outcomes. Research shows that teachers with a strong sense of efficacy enjoy higher levels of job satisfaction (Caprara et al., 2003; Caprara et al., 2006; Tschannen-Moran & Hoy, 2002), have stronger commitment to teaching (Evans & Tribble, 1986; Ware & Kitsantas, 2007), and are less vulnerable to burnout (Brouwers & Tomic, 2000; Shaalvik & Shaalvik, 2007). In addition, efficacious teachers create a better learning atmosphere for their students (Deemer, 2004; Gencer & Cakiroglu, 2007; Woolfolk & Hoy, 1990; Yost, 2002), and are more innovative in the application of new teaching methods (Wertheim & Leyser, 2002). Greater efficacy also helps teachers use their class time as best as they can (Gibson & Dembo, 1984), set high standards for themselves and persist in the face of obstacles (Ross & Bruce, 2007), foster stronger collegial ties (Friedman & Kass, 2002), and enhance students’ achievement (Caprara et al., 2006; Herman, 2000; Midgley, et al., 1989; Ross, 1992; Shaughnessy, 2004; Tournaki & Podell, 2005; Wallik, 2002). So, it seems that the stronger a teachers’ sense of efficacy, the more qualified s/he will be.

On the other hand, two of the most important criteria used by recruiters to
employ second language teachers are professional experience and educational background (Clark & Paran, 2007). However, not much published research is available to show the influence of these features on improving teachers’ quality. The necessity for such a study is seriously felt when one realizes that many of those who are involved in the English teaching profession possess academic degrees in irrelevant majors such as geography, physics, engineering, etc. The present study is a partial attempt to fill this gap with respect to teachers’ efficacy beliefs. More precisely, the relationship between English teachers’ academic degree/professional experience and their sense of efficacy will be addressed. Therefore, the following research questions are posed:

Is there any difference between the self-efficacy beliefs of novice and experienced teachers?

Is there any difference between the self-efficacy beliefs of teachers who have an English-related degree and those who hold irrelevant degrees?

Review of the literature

The concept of self-efficacy (or simply efficacy) was first proposed by Bandura (1977) within his social cognitive theory and refers to “beliefs in one’s capabilities to organize and execute the courses of action required to manage prospective situations” (Bandura, 1995, p.2). Self-efficacy beliefs do not necessarily reflect people’s actual ability, but show their perception of it. Therefore, they may underestimate or overestimate their real abilities (Tschannen-Moran & Hoy, 2007).

Four sources have been postulated to be influential in shaping self-efficacy beliefs (Bandura, 1977, 1995, 1998). The first one is enactive or mastery experience. It is the most powerful source of efficacy and is connected with people’s success or failure in doing a task. “Successes build a robust belief in one’s personal efficacy. Failures undermine it, especially if failures occur before a sense of efficacy is firmly established” (Bandura, 1998, p.53). Vicarious experience, as the second source of efficacy, has to do with the fact that most people try to select models for themselves from among other persons. In such a case, the successes of the chosen model
enhance individuals’ sense of efficacy, especially when there are a lot of similarities between the individual and the selected model. The third source of efficacy is called *social persuasion* which refers to the verbal encouragement people receive from others. If the person who provides verbal persuasion is dependable, individuals’ self-efficacy tends to increase. *Physiological and emotional states* constitute the last source of efficacy and pertain to people’s physical and affective condition during task completion. For instance, feelings of relaxation are signals of self-assurance and, therefore, enhance self-efficacy, while a racing heart beat or high blood pressure can lead to low efficacy beliefs.

When it comes to the academic setting, teacher self-efficacy refers to teachers’ judgment on their abilities to motivate students and improve their achievement (Campbell, 1996; Chacón, 2005; Cruz & Arias, 2007; Hoy & Spero, 2005; Milner & Hoy, 2003; Ross & Bruce, 2007; Ross et al., 1996; Weately, 2005). There are so many factors which may influence this psychological construct, but they can be classified under two broad categories; contextual and demographic factors.

As for the first category, it is said that teacher self-efficacy is a kind of context-specific construct (Chacón, 2005; Dellinger et al., 2008) and is shaped within a particular environment (Friedman & Kass, 2002; Tschannen-Moran, Hoy, & Hoy, 1998). It is supposed to be affected by such factors as the principal leadership and school climate (Tschannen-Moran & Hoy, 2007). More precisely, if teachers have access to more resources in the school and enjoy the principal’s support, they are more likely to have stronger self-efficacy beliefs (Deemer, 2004; Hoy, & Woolfolk, 1993; Tschanne-Moran & Hoy, 2002). In addition, teachers who receive guidance from their colleagues feel more efficacious, regardless of whether it is in the form of supervision (Chester & Beudin, 1996; Coladarci & Breton, 1997), mentoring, or interdisciplinary teams (Warren & Pyne, 1997). Also, the class size can affect teachers’ sense of efficacy in that they possess stronger efficacy beliefs if they teach larger classes (Lee et al., 1991; Raudenbush et al., 1992). Students’ characteristics might affect teacher efficacy as another contextual factor. For instance, Bejarano (2000) found that students’ gender has no effect on teachers’ perceived efficacy (i.e.,
teachers are equally efficacious in teaching both males and females). Evans and Tribble (1986), Herman (2000), Taimalu and Öim (2005), and Tchannen-Moran and Hoy (2002) discovered that teachers are more likely to be efficacious when they teach younger students. Moreover, focusing on students’ social class, Lee et al. (1991) and Hoy and Spero (2005) came to the conclusion that more efficacious teachers are those that teach students who come from the high socioeconomic levels of the society. To sum up, it might be inferred that the context in which teachers work, including the principal, the colleagues, and the students’ characteristics, can affect their self-efficacy beliefs to a great extent.

The second category (i.e., demographic factors) includes variables such as gender, age, experience, and academic degree. Considering gender, for example, available research indicates that male and female teachers do not differ in their perception of self-efficacy (Gencer & Cakiroglu, 2007; Herman, 2000; Hoy & Woolfolk, 1993; Lee et al., 1991; Taimalu & Öim, 2005; Tschannen-Moran & Hoy, 2002; Tschannen-Moran & Hoy, 2007). The only reviewed study in which a difference was found between male and female teachers’ self-efficacy was that conducted by Raudenbush, et al. (1992). Although in their study female teachers had significantly higher level of efficacy than males, this difference was not that much great.

Teachers’ age is another investigated variable with relation to self-efficacy. While Campbell (1996) claimed that older teachers feel more efficacious, Tschannen-Moran and Hoy (2002) could not find any relationship between them. Perhaps this discrepancy could be explained in the light of another variable which covariates with teachers’ age; that is, teaching experience. Older teachers are normally considered to have more experience.

However, the literature seems murky as one tries to see the relation between teachers’ experience and their efficacy beliefs. Some of the researchers have come to the conclusion that teaching experience has nothing to do with teacher self-efficacy (Bejarano, 2000; Chacón, 2005; Gaith & Shaaban, 1999; Howell, 2006; Lee, et al., 1991; Wallick, 2002). Some others (e.g., Cruz & Arias, 2007; Gaith & Yaghi, 1997;
Hoy & Woolfolk, 1990; Taimalu & Òim, 2005;) intended to find the difference between the efficacy of prospective and inservice teachers. They concluded that as teachers enter the profession and gain more experience, their beliefs in their ability to control disturbing factors outside the classroom context, known as general teaching efficacy (GTE), decreases, whereas their beliefs in their own ability to teach within the classroom context, called personal teaching efficacy (PTE) improves. Cruz and Arias (2007) attributed the higher GTE for prospective teachers to the support they receive from their tutors and also to the distance from real classroom situations. As these teachers enter the classroom and confront “the harsh and rude reality of everyday classroom life” (Veenman, 1984, p.143), they discover that educational system is not the sole source for students’ behavior and they are affected by other environmental factors. As a result, their GTE decreases in the course of time. However, these claims were questioned by Huang et al. (2007) who discovered that both GTE and PTE were higher for experienced teachers. Tschannen-Moran and Hoy (2002, 2007) and Chan (2008) attempted to find the difference between the efficacy of novice and experienced practicing teachers. They found that experienced teachers had significantly higher efficacy than their novice counterparts. Tschannen-Moran and Hoy (2007) tried to explain this difference based on the sources of efficacy. Moreover, they found that verbal persuasion significantly predicted novice teachers’ sense of efficacy because “teachers who are struggling in their early years in their careers tend to lean more heavily on the support of their colleagues” (p.953). Experienced teachers, on the contrary, were more likely to take advantage of the strongest source of efficacy (i.e., mastery experience) since they have passed enough time in the career to experience success in their professional lives. As it seems, there is no general agreement on the relationship between teachers’ experience and their self-efficacy beliefs. One of the reasons might be that different researchers have utilized various instruments to measure teachers’ sense of efficacy (e.g., Gibson & Dembo, 1984; Tschannen-Moran & Hoy, 2001). Moreover, they have followed different statistical procedures like Pearson correlation coefficient (e.g., Chacón, 2005) or parametric tests (e.g.,
Tschannen-Moran & Hoy, 2007). Yet, another reason might be that they have had different cut-off points for dividing novice and experienced teachers (e.g., Chan, 2008; Huang et al., 2007). These problems are exacerbated within the realm of English teaching since it suffers from the scarcity of research in this regard. Therefore, it seems that to be able to talk about the relationship between teaching experience and teachers’ sense of efficacy with certainty more studies are required.

The last reviewed demographic characteristic is teachers’ academic degree. It can be approached from two different perspectives; the level and the relevance of the academic degree. The former refers to whether a teacher has higher or lower academic degree within a subject matter, while the latter addresses the issue if a teacher who teaches a particular subject matter has a degree which is related to it. Considering the level, it seems that as teachers get higher academic degrees or go to graduate schools for further education, their sense of efficacy improves (Campbell, 1996; Cantrell et al., 2003; Hoy & Woolfolk, 1993). In addition, on-the-job training programs appear to enhance teacher self-efficacy (Chacón, 2005; Taimal & Öim, 2007; Tucker et al, 2005; Ross & Bruce, 2007). All the mentioned studies, except Chacón (2005), have taken teachers from different disciplines into account. However, when it comes to the relevance of degree, it looks very hard to find a published research which has addressed the problem. It can be due to the fact that, in the case of most subject matters, the relevance of teachers’ academic degree seems to be a prerequisite for recruiting them. For example, it is almost impossible for a teacher without an academic degree in mathematics to teach this subject matter in higher levels of education. But, this assumption is not met about language teachers; as mentioned earlier, it is possible for every one with a good command of English to make an English teacher.

Methodology

Participants

Four hundred and forty-seven English teachers (96 male and 351 female)
participated in this study. Their teaching experience ranged from 1 to 25 years (mean=3.65, SD=3.33). Following Tschannen-Moran and Hoy (2007) and Chan (2008), we chose three years of teaching experience as the cut-off point for dividing novice and experienced teachers. As a result, 253 of the participants were classified as novice and the rest (194 teachers) were categorized as experienced. From among this population, those teachers who had a Bachelor of Art (BA) in English-related majors (i.e., English translation and English literature) or post graduate degrees in Teaching English as a Foreign Language (TEFL) were considered to have degrees which were relevant to English teaching. Other participants were considered as holding irrelevant degrees. Based on this criterion, 202 teachers had relevant and 245 had irrelevant academic degrees.

**Instrument**

Teacher self-efficacy scale (TSES) developed by Tschannen-Moran and Hoy (2001) was used as a criterion to assess teachers’ self-efficacy beliefs since its validity has been proved in different contexts (Klassen et al., 2009). This instrument consists of 24 questions answered on a 9 point likert scale ranging from 1-nothing to 9-a great deal. TSES includes three subcomponents (efficacy for student engagement, efficacy for instructional strategies, and efficacy for classroom management) each of which is measured through eight questions. According to Tschannen-Moran and Hoy’s (2001) guidelines, the instrument can be used in two different ways: first, by calculating the means of participants’ answers to all 24 questions. This score, which ranges from 1 to 9, is called global efficacy; second, by calculating the means of the answers for each of the three subcomponents separately which yields three scores for each person ranging from 1 to 9; the higher the score a participant gains, the higher his/her self-efficacy beliefs.

**Procedure**

In order to control contextual factors as much as possible, the participants were chosen from a single language institute in Tehran. Five hundred teachers working in
fifteen branches of the institute were contacted and invited to participate in this study. After being assured of the confidentiality of the provided information, they were presented with a two-page survey including a consent letter, some demographic information, and TSES. They were requested to take the survey home, answer the questions, and give it back to the principal of the branch who, in turn, would hand them to the researchers for further analysis. Four hundred and forty seven copies were given back which formed the basis for data analysis.

**Data analysis and results**

The obtained information was fed into SPSS (version 16). The values of global efficacy, efficacy for student engagement, efficacy for instructional strategies, and efficacy for classroom management were calculated for each of the participants and the related descriptive statistics are presented in the following table (Table 1).

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global efficacy</td>
<td>447</td>
<td>7.21</td>
<td>0.89</td>
<td>4.25</td>
<td>8.92</td>
</tr>
<tr>
<td>Efficacy for student engagement</td>
<td>447</td>
<td>6.91</td>
<td>0.99</td>
<td>3.86</td>
<td>8.88</td>
</tr>
<tr>
<td>Efficacy for classroom management</td>
<td>447</td>
<td>7.43</td>
<td>0.95</td>
<td>4.12</td>
<td>9</td>
</tr>
<tr>
<td>Efficacy for instructional strategies</td>
<td>447</td>
<td>7.30</td>
<td>0.96</td>
<td>3.75</td>
<td>9</td>
</tr>
</tbody>
</table>

Cronbach alpha was used to assess the reliability of the instrument, showing reliability indices of 0.93 for global teacher efficacy, 0.79 for student engagement, 0.86 for classroom management, and 0.87 for instructional strategies which can be considered satisfactory.

To answer the research questions, four sets of two-way ANOVAs were conducted. In each of them, teachers’ academic degree and experience were considered as the independent variables both of which had two levels; teachers were
divided into relevant and irrelevant groups based on their academic degree and also into novice and experienced categories based on their teaching experience. In the first two-way ANOVA, teachers’ global sense of efficacy was considered as the dependent variable. Descriptive statistics for participants’ scores on global efficacy are presented in Table 2.

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Experience</th>
<th>Academic degree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Novice</td>
<td>Experienced</td>
</tr>
<tr>
<td>Number</td>
<td>253</td>
<td>194</td>
</tr>
<tr>
<td>Minimum</td>
<td>4.25</td>
<td>5.05</td>
</tr>
<tr>
<td>Maximum</td>
<td>8.92</td>
<td>8.79</td>
</tr>
<tr>
<td>Mean</td>
<td>7.00</td>
<td>7.49</td>
</tr>
<tr>
<td>S.D.</td>
<td>0.89</td>
<td>0.81</td>
</tr>
</tbody>
</table>

The results indicated a significant main effect only for experience ($F_{(1,443)}=30.26, p<0.01$) with an effect size of 0.64, but neither the main effect for academic degree nor the experience $\times$ academic degree interaction effect was statistically significant ($F_{(1,443)}=3.81, p=0.058$ and $F_{(1,443)}=0.03, p=0.85$, respectively). In the following plot (Figure 1), it has been schematically shown that experienced teachers had a significantly higher level of efficacy than their inexperienced counterparts. However, although those teachers who had an English-related academic degree enjoyed a stronger sense of global efficacy compared to their colleagues who had non-English degrees, their difference did not reach the significant level.
The second two-way ANOVA, was intended to find the effect of teachers’ academic degree and experience on their efficacy for student engagement. Descriptive statistics for participants’ scores on this subcomponent are presented in Table 3.

Table 3: Descriptive statistics for participants’ scores on efficacy for student engagement

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Experience</th>
<th>Academic degree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Novice</td>
<td>Experienced</td>
</tr>
<tr>
<td>Number</td>
<td>253</td>
<td>194</td>
</tr>
<tr>
<td>Minimum</td>
<td>4.38</td>
<td>3.86</td>
</tr>
<tr>
<td>Maximum</td>
<td>8.88</td>
<td>8.88</td>
</tr>
<tr>
<td>Mean</td>
<td>6.72</td>
<td>7.17</td>
</tr>
<tr>
<td>S.D.</td>
<td>0.99</td>
<td>0.94</td>
</tr>
</tbody>
</table>

The results showed that both the main effects for academic degree and experience were significant ($F_{(1,443)}=5$, $p<0.05$ and $F_{(1,443)}=20$, $p<0.01$, respectively).
with an effect size of 0.11 for the former and 0.43 for the latter. But, the academic degree × experience interaction effect was not significant (F(1,443)=0.004, p=0.95). As the profile plot (Figure 2) illustrates, experienced teachers felt more efficacious than their novice colleagues. Moreover, those teachers who had English-related degrees enjoyed a significantly higher level of efficacy for student engagement compared to those who had non-English-related degrees.

![Figure 2: English teachers' academic degree and experience and their sense of efficacy for student engagement](image)

In the third two-way ANOVA, the effect of teachers’ academic degree and experience on their sense of efficacy for classroom management was investigated. Descriptive statistics of participants’ scores for this subcomponent are presented in Table 4.
Table 4: Descriptive statistics for participants’ scores on efficacy for classroom management

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Experience</th>
<th>Academic degree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Novice</td>
<td>Experienced</td>
</tr>
<tr>
<td>Number</td>
<td>253</td>
<td>194</td>
</tr>
<tr>
<td>Maximum</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Mean</td>
<td>7.22</td>
<td>7.70</td>
</tr>
<tr>
<td>S.D.</td>
<td>0.97</td>
<td>0.87</td>
</tr>
</tbody>
</table>

The results revealed that while the main effect of experience was significant ($F_{(1,443)}=27.37, p<0.01$), neither the main effect of academic degree nor the academic degree × experience interaction effect were significant ($F_{(1,443)}=1.58, p=0.20$ and $F_{(1,443)}=0.10, p=0.74$, respectively). The following plot (Figure 3) shows the same pattern as that of global efficacy; English teachers with more experience had a significantly higher level of efficacy for classroom management, while those who had English-related degrees did not enjoy a significantly stronger sense of efficacy in this regard.

![Figure 3: English teachers’ academic degree and experience and their sense efficacy for classroom management](image-url)
Finally, in the last two-way ANOVA, teachers’ efficacy for instructional strategies was used as the dependent variable. Descriptive statistics for participants’ scores on this subcomponent are presented in Table 5.

Table 5: Descriptive statistics for participants’ scores on efficacy for instructional strategies

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Experience</th>
<th>Academic degree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Novice</td>
<td>Experienced</td>
</tr>
<tr>
<td>Number</td>
<td>253</td>
<td>194</td>
</tr>
<tr>
<td>Minimum</td>
<td>3.75</td>
<td>5.14</td>
</tr>
<tr>
<td>Maximum</td>
<td>8.88</td>
<td>9</td>
</tr>
<tr>
<td>Mean</td>
<td>7.08</td>
<td>7.59</td>
</tr>
<tr>
<td>S.D.</td>
<td>0.97</td>
<td>0.86</td>
</tr>
</tbody>
</table>

The results indicated that neither the main effect of academic degree nor the academic degree × experience interaction effect were significant ($F_{(1,443)}=3.47$, $p=0.06$ and $F_{(1,443)}=0.50$, $p=0.47$, respectively), whereas teaching experience had a significant main effect ($F_{(1,443)}=28.69$, $p<0.01$). Also, as the following plot shows (Figure 4), experienced teachers had a significantly higher level of efficacy for instructional strategies. But, teachers with English-related academic degrees did not have significantly stronger sense of efficacy in this regard.

In brief, teachers who had English-related academic degrees showed a consistent pattern of having a stronger sense of efficacy than those who possessed non-English-related degrees, yet the difference between the two groups was statistically significant only in the case of efficacy for student engagement. In contrast, compared to their novice counterparts, experienced teachers enjoyed a significantly stronger sense of efficacy in all the subcomponents as well as the global aspect.
Discussion

In this section, the results of the study are discussed in relation to the research questions presented earlier.

Research Question 1: The relationship between teachers’ experience and their self-efficacy

The findings are almost in line with Tschannen-Moran and Hoy’s (2002, 2007) who concluded that experienced teachers (with more than three years of teaching experience) have a stronger sense of global efficacy, efficacy for classroom management, and efficacy for instructional strategies compared to their novice counterparts. This difference in the strength of efficacy beliefs among novice and experienced teachers might be attributed to various sources of efficacy they utilize. According to Tschannen-Moran and Hoy (2007), novice teachers, who are struggling to find their voice as they enter the profession, rely more heavily on the support they receive from their colleagues. Consequently, the most noticeable source of efficacy they resort to is verbal persuasion. Experienced teachers, in contrast, tend to lean more toward the strongest source of efficacy (i.e., mastery experience) which
they have accumulated over years. These successful experiences contribute to strengthening the teachers’ sense of efficacy in a cyclical nature, in that, when they succeed in accomplishing a task, they gain greater efficacy which leads to greater efforts and persistence. This, in turn, results in teachers’ improved performance which boosts efficacy (Tschannen-Moran & Hoy, 1998). Therefore, since mastery experience is the most powerful source of efficacy, it enhances experienced teachers’ self-efficacy beliefs to a larger extent in comparison with verbal persuasion which is heavily utilized by novice teachers.

The only discrepancy between the present study and those conducted by Tschannen-Moran and Hoy (2002, 2007) is that while they did not find any significant difference between novice and experienced teachers’ efficacy for student engagement, the present research demonstrated that experienced English teachers feel more efficacious for involving students in class activities. This might be attributed to the difference between the participants in Tschannen-Moran and Hoy’s (2002, 2007) study and those in the present one which was just concerned with English teachers (as opposed to the former ones that chose teachers from different subject matters). Perhaps one of the major differences between teachers of other disciplines and language teachers pertains to the importance the latter group attaches to eliciting students’ engagement in classroom activities. The value of student engagement in language classroom has been echoed by recent theories of language teaching such as communicative language teaching (Savignon, 1991), task-based language teaching (Skehan, 1998), and sociocultural theory of learning (Ellis, 2003). Therefore, a general characteristic of an English teacher, which distinguishes him/her from teachers of other subject matters, is his/her ability to engage students in performing classroom tasks. More precisely, one of the crucial factors which makes English teachers judge themselves as qualified persons is their belief in their ability to elicit students’ participation in the classroom activities. As a result, this subcomponent of efficacy (i.e., efficacy for student engagement) is a distinguishing parameter between novice and experienced teachers.

Another reason for lower levels of efficacy among inexperienced teachers is that while experienced teachers utilize their accomplishments to enhance their sense of
efficacy, their novice colleagues have to overcome the “reality shock” (Veenman, 1984, p. 143) they experience as they enter the classroom for the first time. The result is that their efficacy drops at the early stages of their career (Fives et al., in press; Hoy & Spero, 2005; Rushton, 2000) and it takes time for them to establish a firm and stable sense of efficacy (Tschannen-Moran & Hoy, 1998).

Finally, another factor which might contribute to experienced teachers’ stronger sense of efficacy is the larger number of on-the-job training programs they have attended; research shows that participating in training programs throughout the teaching service can enhance teachers’ sense of efficacy (Tucker et al., 2005), especially their efficacy for instructional strategies and student engagement (Chacón, 2005).

**Research Question 2: The relationship between teachers’ academic degree and their self-efficacy**

The only subcomponent of teacher efficacy in which the main effect of academic degree was significant was the one for student engagement. Even in this case the effect size was not big enough (it was 0.11) to explain a large portion of the variance (Hatch & Lazaraton, 1991). In order to provide a sound explanation, a closer look should be taken at the nature of self-efficacy. As Dellinger et al. (2008) claim, self-efficacy is a context-specific concept and is shaped within a particular environment. Therefore, “[o]nly in situations of actual teaching can an individual assess the capabilities he or she brings to the task and experience the consequences of those capabilities” (Tschannen-Moran et al., 1998, p. 19). So, self-efficacy beliefs are largely behavior-oriented (Bandura, 1977) and if academic programs are intended to increase preservice teachers’ sense of efficacy, they must provide “opportunities for actual experience with instructing and managing children in a variety of contexts with increasing levels of complexity and challenge to provide mastery experience and specific feedback” (Tschannen-Moran et al., 1998, p. 24). Such opportunities might be created through practice-oriented courses rather than theoretical ones. For example, practicum courses which provide a scaffold for student teachers to face increasingly challenging situations in real classroom can
enhance their sense of efficacy (Atay, 2007). Furthermore, to be able to surmount bumpy moments in teaching, efficacious teachers must resort to their personal practical knowledge rather than their theoretical one (Romano, 2006). However, a close inspection of English-related academic programs (English translation, English literature, and TEFL) in Iran reveals that students pass a few courses directly related to the practical side of teaching. Even in those courses (e.g., methodology, teaching skills, and testing), they are usually presented with broad, theoretical underpinnings. Consequently, the share of practice had become less in academic English programs. This is not to say that the present academic programs have nothing to offer with respect to student teachers’ sense of efficacy since, as the results of the study show, there is a consistent pattern of stronger (though not statistically significant) efficacy expectations among teachers who hold relevant academic degrees. But, if these programs are going to play a more influential role in educating efficacious teachers, they have to strike a balance between the theoretical side of the coin, which greatly contributes to teachers’ theoretical knowledge base, and its practical side, which is mostly effective in enlarging their practical knowledge base (Romano, 2006).

Conclusion

The present study was designed to see whether English teachers’ experience and academic degree can make a difference in affecting their sense of efficacy. The findings revealed that while experience was a distinguishing factor, academic degree was not. With respect to experience, the study showed that seasoned teachers with more than three years of teaching experience had significantly higher levels of efficacy than their novice colleagues. This might be attributed to a number of factors. First, whereas experienced teachers take advantage of mastery experience as the strongest source of efficacy, novices are more dependent on verbal persuasion. Second, as teachers enter the profession, they face a reality shock which destroys the dream home they have built for themselves and, consequently, as novice teachers, their sense of efficacy undermines. Finally, experienced teachers have attended a larger number of on-the-job training programs; such programs are shown to have an
influential role in enhancing their sense of efficacy.

Considering the second independent variable, which was academic degree, the findings demonstrated that although teachers with English-related academic degrees showed a consistent pattern of having a stronger sense of efficacy compared to their colleagues with non-English-related degrees (e.g., physics, geography, engineering, etc.), just in the case of efficacy for student engagement the difference reached a slightly significant level with a small effect size. This is largely due to the nature of English-related academic programs and self-efficacy as a psychological trait; whereas self-efficacy is mostly behavior-oriented and context-specific and is shaped as teachers are practically involved in classroom activities, English-related academic programs usually focus on the theoretical aspects of language teaching which have little to offer to increase prospective teachers’ efficacy expectations.

References


---, (2002). "From adolescent angst to adulthood: Substantive implications and measurement dilemmas in the development of teacher efficacy research". *Educational Psychologist*. 37(3): 137-150.


The Long-Term Effect of Selective Written Grammar Feedback on EFL Learners' Acquisition of Articles

Ahsan Pashazadeh*
MA Student, English Department, University of Tehran, I.R. Iran

Hamideh Marefat**
Associate Professor, English Department, University of Tehran, I.R. Iran

(Received: 12 Feb. 2009, Accepted: 10 Aug. 2009)

Abstract

The long-term efficacy of written grammar feedback (WGF) has escaped rigorous empirical investigation, partly due to its strong intuitive appeal. The few studies that have investigated the long-term effect of grammar correction have failed to find any positive long-term effects for WGF (Truscott, 2007). Nevertheless, a number of recent studies have reported a positive effect for WGF. The present study investigated the long-term effect of selective grammar feedback on a reasonably complex feature of the English grammar, attempting to shed some light on the factors that may explain the conflicting results of previous studies. A group of low intermediate EFL learners (N = 22) participated in this study. They received WGF on the use of articles and a limited number of various other grammatical categories. The results showed that selective WGF can produce large short-term gains for functionally complex grammatical features, but that it may prove to be detrimental in the long run. The study also found that learners tend to avoid the grammatical feature on which they have received corrective feedback. The theoretical and practical implications are discussed.

Key Words: Grammar Feedback, Accuracy, Item Learning, System Learning, Articles.
1. Introduction

Several decades of research on grammar feedback have barely yielded any satisfactory evidence regarding the long-term efficacy of providing second language writers with corrective feedback (for reviews of the related studies see Hyland & Hyland, 2006; Norris & Ortega, 2000; Truscott, 1996). The strong claims concerning the positive effects of corrective feedback may have great intuitive appeal; however, these claims have been rebutted on the grounds that the primary studies carried out so far, which have found a positive effect for WGF, have not been designed well enough to provide convincing evidence for grammar feedback.

Believing that feedback is essential to help students recognize their linguistic shortcomings, language teachers spend a great deal of time offering corrective feedback on the written products of their students (Ferris, 1999, 2003; Lee, 2004; Truscott, 1996; Zamel, 1985). Writing teachers' unfailing faith in the efficacy of corrective feedback is shared by the majority of ESL theoreticians, a belief fuelled by the surge of the cognitive accounts of second language acquisition. Researchers working within the cognitive framework invoke the noticing hypothesis and the role of conscious learning to claim that short-term gains in accuracy can be taken as a sign of noticing, which is interpreted as the beginning of learning. Nevertheless, short-term gains may not translate into long-term grammatical accuracy.

In 1996 Truscott published a controversial critical review of feedback studies in which he cited evidence from various studies (such as Frantzen, 1995; Kadia, 1988; Kepner, 1991; Lalande, 1982; Robb, Ross, & Shortreed, 1986; Semke, 1984; Sheppard, 1992; Zamel, 1985) and concluded that feedback is not, and should not be expected to be, effective in the long term. Several researchers (Chandler, 2003; Ferris, 1999, 2003; Ferris & Roberts 2001, to name but a few) criticised him for his strong claims. However, Truscott has not withdrawn his case against grammar correction. In several subsequent papers, Truscott (1999, 2004, 2007) has made it clear that feedback studies do not support any positive effects for grammar correction.

Several years later, the debate still continues between the advocates and
opponents of corrective feedback. In fact, since the publication of Truscott's review article a number of other studies have found corrective feedback to be ineffective (e.g., Fazio, 2001; Polio, Fleck, & Leder, 1998). Nevertheless, Bitchener (2008), Bitchener and Knoch (2008), Ellis, Sheen, Murakami, and Takashima (2008), Rahimi (2009), and Sheen (2007) have provided some evidence that corrective feedback can be effective in improving the accuracy of L2 writers.

A number of feedback studies have investigated the effect of corrective feedback on several grammatical and non-grammatical categories (e.g., Ferris, 2006). However, as Norris and Ortega (2000) have pointed out, studies that try to handle too many variables are not likely to add much to our knowledge base. In the same vein, the effects of error correction have sometimes been measured by grammaticality judgement tests or multiple-choice grammar tests, which clearly favour students who have received formal instruction (see Ellis, 2005). As far as L2 writing is concerned, writing original essays is deemed to be the most authentic measure of writing abilities.

Almost all feedback studies suffer from similar shortcomings – namely, lack of truly equivalent groups (e.g., Bitchener, Young, & Cameron, 2005), inappropriate definition of error categories (e.g., Ferris & Roberts, 2001), lack of a pre-test to measure initial differences (e.g., Kepner, 1991), lack of a delayed post-test to measure long-term effects (e.g., Fathman & Whalley, 1990), and use of inappropriate testing devices (e.g., Cardelle & Corno, 1981).

Ellis et al. (2008) studied the differential effects of focused and unfocused WGF, including a control group to investigate Truscott's (1996) thesis regarding the inefficacy of corrective feedback. The results of the picture story tasks revealed that providing the learners with WGF helped them to improve their use of two functions of the English article system – namely, first- and second-mention use of articles to refer to unknown and known entities. The experimental groups outperformed the control group and maintained their level of accuracy after 4 weeks.

The present study was inspired by Ellis et al.’s (2008) research; however, there were a number of shortcomings in Ellis et al.’s study which the present research has
attempted to eliminate. First of all, the elicitation device used in this study was a picture story task, a measure that could understandably have introduced a memory factor into the students’ writing. In other words, the participants were given a set of pictures (as well as a constructed story on separate sheets of paper, which were collected before the students started writing their narratives) and were instructed to rewrite "the story with as much detail as they could remember" (p. 358). In the present study, controversial topics were given to the students, asking them to express their own ideas on the topics. This measure was taken in order to enhance the validity of the writing tasks.

Second, Ellis et al.’s study used various elicitation devices, including a grammaticality judgement test as well as a questionnaire. These devices were used to triangulate the results. However, the use of a measure like a grammaticality judgement test may give away the purpose of the study. Ellis et al. failed to detect any significant differences between their focused and unfocused treatment groups. Probably, the two treatment groups had similar performances due to the use of the grammaticality judgement test, which made both groups equally conscious of the aim of the study.

Third, Ellis et al.’s study used only one delayed post-test to assess the accuracy of the participants’ writing. Given the U-shaped course of interlanguage development, it is befitting to use a second post-test to better trace the possible long-term effects of the treatment. The English article system is too complex to be significantly affected in a short period of time (Bitchener, 2008; Butler, 2002; Master, 1997). From a cognitive perspective, formal instruction reveals its effects through a process called "restructuring" (see McLaughlin, 1990). This cognitive process is believed to take place as a delayed result of instruction and negative evidence.

Fourth, the study did not account for the overuse of articles. When learners are taught a rule, they tend to overgeneralise it (or avoid the structure altogether). The results of feedback studies have generally been based on the calculation of obligatory occasion analysis, which does not take overuse into account. Sheen
(2007) is an exception in this regard. Obligatory occasion analysis, therefore, leads to overestimation of the accuracy index. Target like use is a measure that can help reduce the possibility of this overgeneralization pitfall (Ellis & Barkhuizen, 2005).

As Ferris (2004) maintains, due to lack of well-designed studies to address the efficacy (or inefficacy) of WGF, "we are virtually at Square One" (p. 56). The present study is a longitudinal attempt to shed some light on the role of grammar correction in helping EFL writers to improve their writing accuracy. The scope of this discussion will be limited to the role of teacher written feedback mainly because it is the major type of feedback offered in EFL contexts.

2. Purpose of the Study

Studies are normally designed to provide an answer to a particular question of some theoretical importance. Before the publication of Truscott's (1996) controversial article in Language Learning, few researchers questioned the role of grammar feedback. Therefore, there are a limited number of studies that have addressed the questions raised by Truscott in a direct and careful manner.

In fact, all the recent studies (i.e., Bitchener, 2008; Bitchener & Knoch, 2008; Sheen, 2007) that have found a significant positive effect for grammar correction in the long term have focused on the use of first- and second mention functions of the English article system. Given the fact that the focus of these studies was formally and functionally simple (Liu & Gleason, 2002; Truscott, 2001), it is difficult to see how the results contradict the thesis advanced by Truscott (1996).

The present research was an attempt to address the shortcomings of Ellis et al. and other recent WGF studies. Specifically, the current study was conducted to investigate the long-term effects of providing grammar feedback on students' original writing\(^1\). A second post-test was included to help better trace the U-shaped course of interlanguage development. Moreover, the entire article system of English was selected to guard against the possibility that formal and functional simplicity of the targeted error categories may render the results theoretically and practically insignificant.
3. Research Questions

This study was designed to answer the following questions:

1- Does WGF help EFL writers to produce more accurate forms of the treated categories?

2- Does WGF produce long-term effects?

3- Does provision of WGF affect avoidance of the treated grammatical category?

4. Methodology

4.1. Participants

This study was carried out in a private language school in Tehran. 22 graduate and undergraduate students – 13 male and 9 male – from an intact low intermediate class participated in this study. They were enrolled in a "conversation class" taught by the researcher. The students met two times a week for 105 minutes. The course had a writing component, although the main focus was on developing conversation skills. The age of the participants ranged from 19 to 34, with the mean age of 23. Out of 22 students, only 15 completed all the writing tasks. Students with incomplete data sets were excluded from the final data analysis.

4.2. Instrumentation

The instruments used for this study consisted of four argumentative writing tasks. Care was taken not to include any topic that would dissuade the use of articles. Moreover, the rubrics were carefully written to encourage the students to use articles in their writing. However, it should be noted that it is very difficult, if not impossible, to force students to use certain structures. Even if it were possible to design such a task, it would not be very desirable, as such a task would pose serious threats to the originality of students' writing.

Moreover, due to the nature of the research questions and practical limitations, it was not possible to include several measures of accuracy. It would have been highly
desirable if the results of the writing tasks had been triangulated through other measures of linguistic accuracy. However, it should be noted that the inclusion of such measures would have introduced other confounding variables into the design of the study.

4.3. Procedure

In the second session of the course, the students were given a reading text on a controversial topic and were instructed to read it in approximately 5 minutes. It was explained to the students that they would be required to write their own ideas about the subject introduced in the reading text. The students were allowed to ask questions in case they had difficulty understanding the reading passage. The passages were constructed in a careful manner to convince the students that it was possible to argue without using low frequency words and complex grammatical structures. After the reading passages had been collected, the students received the writing task instructions at the top of a separate sheet of paper. The students were required to write an argumentative essay in 25-30 minutes on the same topic as that of the reading text.

The essays were photocopied and subjected to data analysis, which yielded the pre-test scores. The original essays were corrected and returned to the students the next session. The participants were provided with the correct forms as well as a general content comment at the end of their essays and were asked to take a look at the form and content comments for about 5 minutes. Afterwards, they received the second reading passage, and the same procedure was adopted for the next two essays. The third essay was analysed for accuracy, the results of which are reported as the immediate post-test. Another task (i.e., delayed post-test 1) was administered with a four-week interval. The students were not told when they would be required to write another essay so that they could not prepare themselves for the writing task. As the second delayed post-test, the same pre-test writing task was administered four weeks after the first delayed post-test (the significance of this measure will be explained below). On the two delayed post-tests, the students received neither
content nor grammar feedback. The timetable is displayed in Table 1.

<table>
<thead>
<tr>
<th>Session</th>
<th>Testing and feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 (week 1)</td>
<td>The first writing task (i.e., pre-test)</td>
</tr>
<tr>
<td>3 (week 1)</td>
<td>feedback on the first writing task; the second writing task</td>
</tr>
<tr>
<td>4 (week 2)</td>
<td>feedback on the second writing task; the third writing task (i.e., immediate post-test)</td>
</tr>
<tr>
<td>5 (week 2)</td>
<td>feedback on the third writing task</td>
</tr>
<tr>
<td>13 (week 6)</td>
<td>The fourth writing task (i.e., delayed post-test 1)</td>
</tr>
<tr>
<td>21 (week 10)</td>
<td>the fifth (same as the pre-test task in terms of content and instructions) writing task (i.e., delayed post-test 2)</td>
</tr>
</tbody>
</table>

The students received up to seven corrections on each essay, two of which were always article errors. Of course, it must be noted that some essays did not contain two article errors or seven errors on the whole, in which case as many errors as possible were corrected. The participants also received a single general comment on the content of their writing – such as 'Good' or 'This is really persuasive' or 'Pay more attention to the ideas'. Moreover, care was taken not to provide any oral or written feedback on the use of articles during the two four-week intervals.

4.4 Scoring

Target like use (see Sheen, 2007) was utilized in order to answer the first and second research questions (the number of articles used correctly divided by all the contexts in which the target language would require a particular article plus the number of articles oversupplied in non-obligatory contexts). Based on this analysis, each essay was assigned a score, which was calculated as a percentage. These scores were fed into SPSS, and paired samples t-tests were run to see whether there were any significant differences at different testing points.
In order to answer the third research question, obligatory occasions were identified on the written products of the students (i.e., all the contexts in which the target language would require a particular article). If feedback causes learners to avoid the treated categories, then learners should avoid the contexts that necessitate the use of the treated grammar category. Therefore, the number of obligatory occasions can be used to shed some light on whether and to what extent the participants tend to avoid the target grammatical category (Ellis et al., 2008).

All the essays from the four testing times were scored collaboratively with an MA student of TEFL. The essays were corrected independently by the two raters. Having discussed the points of disagreement, perfect agreement (i.e., 100%) was reached. Sometimes, due to numerous grammatical and lexical errors, it was not clear whether or not an article should have been used. Therefore, the contexts that did not provide any clear clue about the use of the correct articles were left out of the final analysis. This decision was made in order to reduce the effect of subjective scoring. All the 15 essays from the pre-test were rescored by the author 14 weeks after the initial scoring, and the Pearson Product Moment Correlation yielded a reliability score of .91.

5. Results and Discussion

The descriptive statistics from the four testing times are displayed in Table 2 and graphically illustrated in Figure 1 below. Table 2 shows that the mean accuracy scores increased substantially from time 1 to time 2 (i.e., from pre-test to immediate post-test), only to decline drastically at the two delayed testing points (i.e., delayed post-tests 1 and 2). The paired samples t-test revealed that there was a statistically significant difference between pre-test and immediate post-test ($t_{(14)} = 4.163; p = .001$). As has been demonstrated by other feedback studies, this finding corroborates the observation that corrective feedback is effective in the short run.
Table 2: Descriptive statistics: accuracy over time

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>15</td>
<td>35.64</td>
<td>12.40</td>
</tr>
<tr>
<td>Immediate post-test</td>
<td>15</td>
<td>53.86</td>
<td>17.95</td>
</tr>
<tr>
<td>Delayed post-test 1</td>
<td>15</td>
<td>38.66</td>
<td>17.31</td>
</tr>
<tr>
<td>Delayed post-test 2</td>
<td>15</td>
<td>32.60</td>
<td>19.18</td>
</tr>
</tbody>
</table>

The fact that corrective grammar feedback is effective in the short run, however, does not guarantee that it will be effective in the long run. Two more paired samples t-tests were conducted to reveal the differences between pre-test and delayed post-test 1 on the one hand and between pre-test and delayed post-test 2 on the other. The results of these tests revealed that there were no significant differences between pre-test and delayed post-test 1 ($t_{14} = 0.698; p = .497$) and between pre-test and delayed post-test 2 ($t_{14} = 0.541; p = .597$). This finding lends further credence to Truscott's (1996) "correction-free approach" (p. 116).
As pointed out by Norris and Ortega (2000) and Truscott (2007), tests of statistical significance do not show magnitude of the differences observed among different variables. Thus, Cohen's $d$ was selected to test the effectiveness of the treatment in the present study. The results are presented as effect sizes in Table 3. Following Cohen (1992), the value of $d$ will be interpreted as follows: small effect $= 0.20$ to $0.50$, medium effect $= 0.50$ to $0.80$, and large effect $= 0.80$ and higher.

Table 3: Effect sizes

<table>
<thead>
<tr>
<th>Pre-test and immediate post-test</th>
<th>Pre-test and delayed post-test 1</th>
<th>Pre-test and delayed post-test 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d = 1.18$</td>
<td>$d = 0.20$</td>
<td>$d = -0.19$</td>
</tr>
</tbody>
</table>

According to Table 3, there is a large effect size for feedback on the immediate post-test, which almost disappears on the first delayed post-test. The effect size for delayed post-test 1 may be interpreted as evidence for grammar correction; however, it should be noted that this effect size is not substantial enough to support this interpretation. In fact, Norris and Ortega (2000) reported that the control groups in their meta-analysis displayed similar gains: $d = 0.30$. Thus, the gains in accuracy can be attributed to writing practice (see Rahimi, 2009), maturation, and other factors.

Table 3 also shows that there was a very small negative effect for feedback on the second delayed post-test. This finding is consistent with the results of some previous corrective feedback studies (e.g., Fazio, 2001; Sheppard, 1992). Corrective feedback – and instruction in general – causes drastic improvement in accuracy levels of learners in the short run. Nevertheless, as Krashen (1981) and Truscott (1996) argue, these changes are caused by learners' metalinguistic knowledge and conscious control over their output. These superficial changes do not reach the learners' competence, which may be one reason why they tend to disappear over time.

Considering the fact that the second post-test was the same as the pre-test
writing task in terms of the reading passage and the instructions, the negative effect size on the second post-test should be particularly alarming to the proponents of WGF. There is some evidence to suggest that learners cannot transfer instruction to other contexts (James, 2008). That is to say, task similarity (or difference) is an important factor that can affect learners' performance. For instance, instruction may help learners to use the English articles accurately in the narrative mode, but they may fail to use them accurately in another task that involves argumentative writing.

Ellis (1999) distinguished between system learning and item learning. He suggested that learners learn through both systematic internalisation of rules (system learning) and memorization of isolated items (item learning). The English article system is functionally complicated (Butler, 2002; Liu & Gleason, 2002; Master, 1997). It may be argued that the brief treatment in this study could not have contributed to the acquisition of the article system, which demonstrates several idiosyncratic, culture-based characteristics. In fact, the second delayed post-test was administered with this in mind. Assuming that grammar feedback is effective, learners should be able to avoid making the same mistakes on an identical task. The findings of this study are significant in that they fail to support any long-term effect for both system and item learning.

It has been argued that learners tend to avoid the categories that have been the subject of corrective feedback (Sheppard, 1992; Truscott, 1996, 2004). Semke (1984) also found that students who received correction simplified their essays in order to avoid being corrected. Truscott (1996) persuasively argued that this avoidance strategy thwarts natural language acquisition simply because grammar feedback encourages students to avoid risk-taking and experimenting with the grammatical forms that they have not fully mastered.

The third research question in the present study was formulated to provide some evidence on whether or not students avoid the treated grammatical categories. To this end, the obligatory occasions were counted. Following Ellis et al. (2008) and Truscott (2007), it was assumed that if feedback causes learners to avoid the categories that have been corrected, the number of obligatory occasions should
decrease over time. Table 4 shows the mean obligatory occasions in this study for the four testing times.

Table 4: Mean obligatory occasions

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>15</td>
<td>11.06</td>
<td>8.64</td>
</tr>
<tr>
<td>Immediate post-test</td>
<td>15</td>
<td>8.80</td>
<td>5.21</td>
</tr>
<tr>
<td>Delayed post-test 1</td>
<td>15</td>
<td>11.26</td>
<td>5.96</td>
</tr>
<tr>
<td>Delayed post-test 2</td>
<td>15</td>
<td>10.00</td>
<td>4.72</td>
</tr>
</tbody>
</table>

It should be noted that the important comparison here is between the pre-test and the immediate post-test, as the students received corrective feedback only on their first, second (not represented in the table above), and third essays. According to Table 4, students do tend to avoid contexts that necessitate the use of articles, although the difference did not reach statistical significance ($t_{(14)} = 1.065; p = .305$).

6. Conclusion

The results of the present study are not consistent with the recent feedback studies (e.g., Bitchener, 2008; Bitchener & Knoch, 2008; Ellis et al., 2008; Sheen, 2007). Nevertheless, the inconsistency can be easily explained by recourse to the differences between the elicitation device as well as the target structure of the present study and those of the recent studies mentioned above. It should be noted that all these studies used highly selective grammar feedback and focused on the use of only two functions of the English articles (i.e., first- and second-mention functions to refer to unknown and known entities), employing pictures as the elicitation device. Moreover, except for Sheen (2007), none of these studies took account of the possibility of overuse, which may have affected the results.

Rahimi (2009) investigated the effect of corrective feedback on the accuracy of
EFL writers. He offered extensive corrective feedback on the expository and argumentative essays written by Iranian English majors. Rahimi concluded that the results of his study did not show a significant effect for feedback alone, although he maintained that corrective feedback had revealed its effect in interaction with practice. Therefore, Rahimi's results are fairly consistent with other feedback studies.

In fact, most of the criticisms that have been levelled against Truscott are rooted in a misunderstanding of his 1996 case against grammar correction. For instance, Ferris (1999) has argued that Truscott's sources are out-dated, that he has ignored evidence that run against his thesis, that feedback has been studied in various contexts and cannot be generalised to other contexts, and that he has rushed to an erroneous conclusion. Truscott (1999) published a response to Ferris's criticisms in which he dismissed most of her arguments.

In terms of interlanguage development, Truscott's (1996) thesis is theoretically more appealing than the rival claims. From an affective point of view, corrective feedback may have disastrous effects on students' attitudes (Hendrickson, 1980; Semke, 1984). Moreover, the assumption behind feedback studies is that there is a linear relationship between feedback and L2 accuracy. However, ESL researchers generally accept that L2 proficiency develops in a complex, U-shaped manner (Kellerman, 1983 cited in McLaughlin, 1990). Therefore, one might wonder why corrective feedback could ever be effective. Feedback studies are based on the positive role of consciousness in learning and on the noticing hypothesis (see Schmidt, 1990, 1994). According to this hypothesis, learners' conscious attention is a necessary and sufficient condition for acquisition to occur. Truscott (1998) has argued that the noticing hypothesis lacks a theory of language and that it does not have a firm basis in psychological research.

Sometimes applied linguists and SLA researchers appear to ignore the early findings regarding the course of second language development. Applied linguists are aware of the non-linearity of the acquisition of grammatical rules and regularities. However, they tend to oversimplify language acquisition processes (see Larsen-
The present study has theoretical as well as practical implications. Theoretically, this study contributes to the debate on the role of negative evidence in promoting second language acquisition. From a practical point of view, this study may help dispel the misconceptions regarding the positive effects of (extensive) corrective feedback, which is a time-consuming practice for teachers. Teachers spend a great deal of time giving formal feedback on the written products of their students – in fact, busy teachers may provide inconsistent and careless feedback due to lack of time (Zamel, 1985). However, there are good reasons to believe that students do not benefit from extensive (or selective?) feedback: Negative affective impacts, lack of attention on the part of students, and learners' in-built syllabi (Corder, 1967).

A number of factors, which have traditionally been included in other feedback studies, were excluded from the present study. This measure was taken to prevent the uncontrolled effect of these factors on the final results of the research. First, in order to guard against the consciousness raising nature of formal tests and the possibility that they may blur the differential treatment effects, no such instruments were included in the present study. Second, there is some evidence to suggest that different structures react differently to negative evidence (Ferris, 2006; Rahimi, 2009; Truscott, 2001). Following the recommendation by Norris and Ortega (2000), this study focused on the effects of selective grammar feedback on the development of only one grammatical category.

Like any other research, there were a number of factors that were not addressed by the present study. First and foremost, this study did not have a control group and the results should, therefore, be interpreted with caution. Second, this study did not include a proficiency test to determine the general proficiency of the students, which is a major shortcoming in most feedback studies carried out in the past. Third, it has
been suggested that various personality factors may differentially affect the success of grammar feedback (Guenette, 2007; Sheen, 2007). Nevertheless, due to practical limitations, it was not possible to measure the students' individual characteristics.

Notes

As Truscott and Hsu (2008) have demonstrated, successful incorporation of corrections into subsequent drafts is not a good measure of learning. Thus, revision studies (e.g., Maftoon & Rabiee, 2007; Sayyad Shirabad, 1999) have not been considered in the present article.

The results of a pilot study showed that intermediate (or more advanced) students demonstrate high levels of accuracy on the target structure (i.e., the English article system). Thus, less proficient students were selected for the purpose of this study in order to avoid a "ceiling effect".

References


Classicism and/or Romanticism: A Survey of Aesthetics in

The Winter's Tale

Zahra Jannessari Ladani∗

PhD Student of English Literature, Faculty of Foreign Languages and Literatures, University of Tehran, I.R. Iran

Fazel Asadi Amjad**

Assistant Professor, Faculty of Humanities, University of Tarbiat Moallem, I.R., Iran

(Received: 25 Jun. 2008, Accepted: 20 Oct. 2009)

Abstract

This article aims at reading The Winter's Tale with respect to Hegel's argument in Aesthetics: Lectures on Fine Art. The researchers try to apply Hegel's notion of art to the artistic aspects of the mentioned play by Shakespeare. The play's aesthetic values will be examined and illustrated from two differing but interdependent classical and romantic perspectives based on Hegelian definitions pertaining to diverse epochs through which art and its representations were developed. The researchers will show that the play contains characteristics of both classical and romantic art forms, but this does not necessitate that Shakespeare believed in these notions and exploited them quite innocently. Rather, as the objective of this study will manifest, it will be proved that Shakespeare neither approved of the classical notions of art nor adopted the romantic version, but stood on the borderline in between and manipulated them intelligently and playfully giving a parodic version of both aesthetics.

Key Words: Shakespeare, Hegel, Aesthetics, Classic Art Form, Romantic Art Form, Sculpture, Paradox.

* Tel: 021-61119053, Fax: 021-88634500, E-mail: zjannessari@ut.ac.ir
** Tel: 021-61119053, Fax: 021-88634500, E-mail: asadi@saba.tmu.ac.ir
Introduction

Then sculpture and her sister-arts revive;
Stones leaped to form, and rocks began to live.

(Pope 2524: 701-02)

In "Romanticism and Classicism" T. E. Hulme brings forth the salient and long-debated argument of classicism and romanticism contrasted; in brief, he prophesies a classical revival in future literature. But the thing is that, in his debate, he introduces Racine as an extreme example of a classical poet. Coming to Shakespeare, however, he does not put him on the romantic side, but adopts Nietzsche's categorization of classicism dichotomized in two levels: static and dynamic. Hulme believes that Shakespeare belongs to the "classic of motion" (Hulme 119). But to what extent could we rely on this statement? Or how can we measure Shakespeare's tendencies toward either classicism or romanticism? How can we be sure that he is not treating the two art forms grotesquely or parodically? These questions will be raised as soon as one reads Hulme's claim.

The example of Hulme has just arbitrarily been chosen for the purpose of this essay. Thus, the article is to contend neither Hulme nor other writers who, on the contrary, classify Shakespeare as Romantic. Here, Hulme's judgment about Shakespeare will be put to test to find out to what extent the playwright employed either Classic or Romantic aesthetic notions in his works. To fulfill this aim, on the one hand, the scope of this study will be confined to the aesthetic aspects of *The Winter's Tale*; on the other hand, a Hegelian model of reading through the play's aesthetics will be followed.

Discussion

From a Hegelian perspective, fine art originates from the fact that the idea of the Absolute comes to be known and realized only when it materializes itself in the form of beauty (Hegel 613-14). Now this form, responsible for the representation of the idea, adopts different manifestations in diverse epochs. In the first epoch, the
Symbolic, the identity of content and form are not reached but there is merely a relationship between the two. We have only the indication of the inner meaning in an external appearance similar to that indication and the content it is supposed to manifest (Ibid 624). Put it differently, the sublime and its artistic form superficially agree with each other. That is why the form peculiar to this epoch, namely architecture, cannot represent the Ideal sufficiently; hence its disproportionate and heterogeneous nature.

The second epoch is associated with the Classical art whereby there is perfect harmony between matter and form. In addition, the classical ideal "corresponds to the portrayal of the Absolute" in its "independently self-reposing reality" (Ibid). Here, the idea acquires more spiritual individuality and is more realistic, since it well intermingles with the sensuous and corporeal. The form pertaining to this epoch is sculpture and plastic arts which are more individualized and conscious than the Symbolic form of architecture.

In the third epoch, art attains still more consciousness; romantic art introduces the idea as the sublime and tends toward infinitude so much so that no form (because form is finite) could appropriately reproduce that spirituality. So the form remains foreign and indifferent. Romantic art, thus, turns inward, but inexorably needs some kind of form to express itself. It has for both its "content and form the subjectivity of emotion and feeling in its infinity and its finite particularity" (Ibid). The forms are poetry, painting and music.

Hegel believes that the Symbolic epoch is the initial stage of art production pursuing the aspiration toward the Ideal as the true Idea of beauty. The idea, though the conception of the mind, has to be expressed via sensuous phenomena which are eventually unable to represent it perfectly. Architecture is the form for this stage as the beginning of art. It has not yet found the adequate material or the corresponding forms for the presentation of its spiritual content. The material is still inherently "non-spiritual, i.e. heavy matter, shapeable only according to the laws of gravity" (Ibid). Thus, forms such as monuments, emblems, and architecture are ineluctably paradoxical, ambiguous, enigmatic and mysterious. However, the heterogeneity of
idea and form does not necessarily mean that material and form are arbitrarily juxtaposed; here, form and idea are worked out by imagination to create the beautiful and the free. This stage seems to be purely abstract and unconscious.

Moving on to the Classical epoch, the art form still lays hold of heavy material in its "spatial entirety" with the exclusion of "inorganic" (Ibid 625) aspects. This form, heavily determined by the content, is "the real life of the spirit, the human form and its objective organism, pervaded by spirit" (Ibid). The form should be adequate, and capable of showing a lofty peace and tranquil greatness. So it should not be touched by the "disunion and restriction of action, conflicts, and sufferings" (Ibid). The proper art form possessing these features is sculpture (Ibid 624).

In "Concerning the Relation of the Plastic Arts to Nature", Wilhelm and Schelling enumerate prominent characteristics of the art: that it is wordless and its source is the soul; that it should be expressed like silent nature by the configuration of sensuous works (Wilhelm & Schelling 324). They argue that the plastic works are, furthermore, empty because the sublime and absolute cannot be demonstrated in them; hence, the intensification of relativity (Ibid 328), and multiplicity. Plastic arts are considered to be a moderation of expression of passions due to their observation of two principles: activity (representative of the soul) and rest (representative of matter) (Ibid 348). Compared with the Symbolic art forms, the Classical ones enjoy more consciousness.

In the romantic phase, the object of human consciousness is no longer God as God; here, the meaning of God represents itself as man's "actual life of subjectively living action and suffering, or the spirit of the community, spirit with a sense of itself, mind in its privation, its sacrifice, or its blessedness and joy in life and activity in the midst of the existing world" (Hegel 625). Now the idea is forced out of reconciliation with the corporeal and backs upon itself. Hegel contends that this art form cannot use heavy matter with its three-dimensional character. Rather, it has to "inwardize or spiritualize" (Ibid) that matter. This is done through the cancellation of the real and sensuous appearance.

Hegel recommends painting, poetry, and music as tolerably representative of the
Romantic spirit. Painting brings together the "trinity of spatial dimensions" of sculpture into a surface and presents "spatial intervals and shapes" (Ibid 626) through color. Music yearns for the same spirit in a different way: it is the manifestation of subjectivity and particularization though still sensuous. It excludes space completely. The visibility of painting is converted to the audibility of the sound in music. According to Hegel, the material for music is sound and its configuration is "counterpoint, the harmony, division, linkage, opposition, discord, and modulation of notes in accordance with their quantitative differences from one another and their artistically treated tempo" (Ibid). Music might be considered as the intermediary between painting and poetry. As for poetry or lyricism, the form tends toward the concretization of the idea by retaining the audibility (spoken syllables) and visibility (letters on the page). Yet, as Hegel argues, sound and letter do not possess a value on their own account. They are wholly permeated with the spiritual world and the particular objects of "ideas and contemplation". So they are the mere external designation of the content. Thus, the objectivity of poetry lies in its objective way of presenting something to inner imagination (Ibid 627).

So far, we have given a very concise introduction about Hegelian notions of fine art, its development, and shifts in diverse contexts. What follows is a track-down of the above mentioned Hegelian progressive process from Classicism to Romanticism in *The Winter's Tale*. The Symbolic art forms will not be examined in the play just to limit the scope of the study to the mentioned shift between the two later epochs. The researchers do not wish to maintain that Shakespeare precisely adhered to the Classic-to-Romantic shift in his play, but to discuss the following questions: to what extent did Shakespeare conform to the linear progression from the Classic to the Romantic aesthetic forms in *The Winter's Tale*? Did he approve of one and dispense with the other or did he deploy a combination of both? How did he treat each of the forms or their combination?

To answer the above questions, we will firstly focus on the notion of sculpture as a plastic art form employed by Shakespeare in the final scene where Hermione, Leontes's wife, who has been in hiding for sixteen years, performs the role of a
statue and gradually comes to life as a complete human-being helped by Paulina's make-believe role as a conjuror. The story is not Shakespeare's own invention. It has a precedent in the Ovidian myth of Pygmalion. A greatly gifted sculptor, Pygmalion created the statue of the most perfect woman on the earth and fell in love with her desperately. Eventually, his wish was fulfilled by Venus who inspired life into the statue and graced their marriage by her presence (Hamilton 1959: 108-11). Perhaps, one could confidently say that Shakespeare made use of the Ovidian story, but one would doubt if his attitude toward the story was anything like Ovid's. That is why Shakespeare cannot be declared a totally Classic-minded playwright; hence, the necessity to locate the transmutations he brought to the Classic Ovidian story. Prior to exposing these transmutations, the extent to which Shakespeare embedded Classical notions of plastic arts in his play needs to be elucidated.

With respect to the major principles of artistic creation in Classicism, Horace had said in On the Art of Poetry that the artist's creation should be harmonized, plausible, and creative (Horace 79); furthermore, the artist should embrace a faithful imitation of things (Ibid 83) and keep the dignity of his artwork intact. He should not allow mockery to trample his work's dignity and should, therefore, elude the introduction of trivialities (Ibid 87). The artwork, as a form, should do its best to manifest the Ideal. As Bate argues in From Classic to Romantic Premises of Taste in Eighteenth Century England, the "representation of the human being in classical sculpture … does not … seek to evoke images from the past experiences and thus appeals to the affections … of the beholder, but rather, by an imitation of the ideal, to form and control those affections" (Bate 19). Thus, the representation of deformities or peculiarities is regarded as defective in imitation of the ideal. Here, Bate renders the particular example of Alexander the Great whose short stature was not represented in the works of the great Italian painters of the Renaissance. He elaborates on the fact that in Classical sculpture, as in Classical antiquity, imitation of the ideal nature of man must "include the presentation of a finished figure, which only the fullest physical development, in every respect, could have formed" (Ibid 80).
Proclivity to the ideal, the sublime, and the perfect, despite the existence of real deformities, is closely interwoven with the aesthetic notions inherent in The Winter's Tale. A good example is the third gentleman's account of Hermione's statue yet to be uncovered by Paulina:

[A] piece many years in doing, and now newly perform'd by that rare Italian master, Julio Romano, who, had he himself eternity, and could put breath into his work, would beguile Nature of her custom, so perfectly he is her ape: he so near to Hermione hath done Hermione, that they say one would speak to her, and stand in hope of answer. (V. II. 81-6)

The excerpt evinces the ideal imitation of Hermione, her defects whatsoever being wholly precluded.

Hermione's statue procures a sort of serenity and calmness in addition to its prominent perfection. This matches the spiritual individuality and isolation Hegel had noticed in the Classical art form of sculpture, especially visible in gods' statues. Paulina elaborates on this aspect of Hermione's statue:

As she lived peerless,
So her dead likeness, I do well believe,
Excels whatever yet you lookt upon,
Or heard of man hath done: therefore I keep it
Lonely, apart.
[…]
I like your silence,—it the more shows off
Your wonder…. (V. III. 17-25)

Ideal imitation has so perfectly been done that Leontes deems himself "being more stone than" Hermione's statue. He calls it a "royal piece" asserting that there is "magic" in its "majesty." Here, idealization causes the statue to seem life-like: "masterly done:/The very life seems warm upon her lip" (Ibid 78-9). The statue turns into the sublime, but Shakespeare does not halt there. He moves on with these Classic notions up to the point where a turn of attitude could apparently be
discerned. The statue gradually comes to life which is not something new based on what we mentioned about the myth of Pygmalion; nonetheless, Shakespeare's attitude here makes his version of the story very different from Ovid's. Ovid's story is, indeed, quite seriously narrated and observes the mentioned dignity characteristic of the Classical art forms. In contrast, Shakespeare's mocking treatment of the plastic art form cannot be denied in *The Winter's Tale*. Bestowing breath, sight, warmth and motion on a statue is a sort of resurrection. Still, resurrection per se is the derision of the whole notion of sculpture. Hermione as statue is originally Hermione as human-being playing the role of a stone piece of artwork. In other words, Hermione has never been a real piece of sculpture but just a fake but simultaneously living one. Thus, the sublime, the ideal, and the serene of sculpture are dissolved into the mortal flesh, the real, and the dynamic. This reminisces Bakhtin's distinction between the "classical" and the "grotesque." In *Rabelais and His World*, Bakhtin tracks down a tradition of portraying the body as fluid, open, changing, procreative, unfinished and belonging to a continuous cycle of living and dying that reveres decay, death, dirt, and excrement as vital to life, growth, order and beauty (Bakhtin 23-30). In this play, too, the existence of wrinkles on Hermione's brow implies the grotesquery of the degenerating body in Bakhtin's theory. As the Renaissance proceeds, Bakhtin discovers an increasing tension between the "grotesque" version of the body with its dynamism and the "classical" version where there is stasis. At this point, it may come across the mind that Shakespeare might have intended to critique the whole classical notion of art form. After all, this question always haunts the mind that if Classical art forms are truly ideal and perfect, why should they be so unfaithful in terms of not representing the exact object of imitation, i.e. why should they render them better than they are? And, again, if they are ideal and perfect, why did there appear a vital need to revolutionize Classic art forms by the advent of Romanticism?

Mockery (resulting in grotesque) is what Paulina's sarcastic and playful tone reveals at the moment she draws back the curtain to present the statue:
But here it is: Prepare
To see the life as lively mockt as ever
Still sleep mockt death: behold, and say 'tis well. (V. III. 21-3)

It is strongly paradoxical and derisive that the nonexistent Romano's artwork should be both perfect and deficient. Notwithstanding all the praises which go to the statue, Leontes notices the fact that "Hermione was not so much wrinkled, nothing/ So aged as this seems" (Ibid 32-3). However, this paradox precisely comprises the turning point of the play whereby one could grasp a shift from Classical to Romantic premises of aesthetic form. The shift exhibits itself in the form of the statue's transcendence of sculpture and its venturing into life; hence, the movement from Classicism to Romanticism. It seems that the idea cannot tolerate to be imprisoned within the static form of sculpture and, consequently, opts for a higher, more spiritual and subjective status. Thus, Hermione's resurrection could possibly be viewed as insurrection accompanied with a dead object's stepping over the threshold and joining living beings. We should not leave out the fact that, as Scott Wilson informs us, Shakespeare's time was the reformation age of revolt against the notion of "master" (Wilson 93); hence, his derisive use of the non-existent Julio Romano as the master sculptor. In addition, Michael Hattaway argues that in The Winter's Tale, Shakespeare "satirizes naïve belief in the authority of print" (Hattaway 50). One of the other issues relevant to the age of reformation is iconoclasm discussed by Marion O'Connor in "'Imagine Me, Gentle Spectators': Iconomachy and The Winter's Tale." Here, the writer defines iconoclasm as the "physical defacement or destruction of religious images, particularly three-dimensional ones, crashed through English ecclesiastical buildings in successive waves for more than a century from 1536" (O'Connor 366). The above notions of revolt and rebellion against authority could possibly justify Shakespeare's attempt to do away with Hermione's statue by breathing life to it.

Yet, this higher spirituality and abstraction cannot avoid form without which it is unable to represent its beauty, perfection and excellence. As we know, Hegel
proposed poetry, painting, and music as romantic art forms. In this play, however, it seems that a new Romantic art form is being introduced. Shakespeare seems to suggest "human being" as a possible form; besides, we could also notice the presence of music as an already existing Romantic art form performed precisely at the moment of Hermione's metamorphosis:

Music, awake her; strike!—[Music.
'Tis time; descend; be stone no more; approach;
Strike all that look upon with marvel. Come;
I'll fill your grave up: stir; nay, come away;
Bequeath to death your numbness, for from him
Dear life redeems you.—You perceive she stirs:
[Hermione comes down from the pedestal. (V. III. 121-26)

Nevertheless, Shakespeare's use of Hegelian as well as novel Romantic forms should not convince us that the playwright approved of and believed in the perfect appropriateness or superiority of the Romantic art forms. We should bear in mind that he brought reconciliation between Classical and Romantic aesthetic notions only through their tantalization on several occasions. Moreover, the Shakespearean innovative shift from dead object-hood to living subject-hood could be asserted as one of the major influences of the salient distinction between Natura Naturans (creative nature) and Natura Naturata (created nature). Based on this distinction, Renaissance artists believed they should not imitate the mere phenomena of nature (natura naturata) but, rather, its higher invisible principles (natura naturans), and this way they could successfully emulate Nature. This, they believed, would impart enlightenment to the art viewers. Also, from the thematic point of view, G. Wilson Knight's reading of the play as "great creating Nature" is worth mentioning (Overton 20). Furthermore, the notion of a statue coming to life hints at the act of God's creation of Adam. God's creation, at least in one respect, could imply the abolishment of idolatry. However, the creator in this play is not God, but Shakespeare. More interestingly, he does not create a man but a woman. It seems
that Shakespeare's revolt targets all extant versions of creation, both artistic and biblical.

The theme of marriage between diverse aesthetic forms, art and nature, the ideal and the real, and "double killing" could be detected throughout the play. One of the best instances in this respect belongs to Perdita and Polixenes's discussion about floral interbreeding. He tries to cajole Perdita that such an art which marries a "gentler scion to the wildest stock" and creates "a bark of baser kind/ By bud of nobler race," will add to nature and mend it; "but/ The art itself is nature" after all (IV. IV. 108-12). On a higher level, such a condition could also be associated with Perdita and Florizel's marriage; the former being a peasant girl, the latter a prince.

Both Perdita and her mother, Hermione, are doubly killed and created. Both were reduced to nothingness and then again discovered and rewarded. Being born, Perdita was cast into the wilderness with a bundle of relics. She was literally denied and left to nothingness, death, and nihilism. Even later on, with the old shepherd, her real identity was hidden from her; so she was masked, doubly negated, and brought up with a fake identity. Afterwards, she was discovered and made princess. As a result, the biased banishment imposed by Polixenes against his son's marrying her dissolved of its own accord. Howard Caygill labels the process as "double negation": "nothing is understood as the negation of being which is itself negated in a turn to being" (Caygill 106). This statement also perfectly goes with Hermione's
resurrection. She was convicted of lechery and high treason, and subsequently ostracized; this led to her virtual death. Then, after sixteen years of muteness, absence, and non-existence, she was found to be living. This did not take place normally though. Being forced to play a dead statue's role, she went through her second negation. This process of "double negation" might as well be displayed by means of the following diagram (the loop happens once in the play, but since the notions of negation and resurrection are ubiquitous, one might view the diagram as a vicious circle too):

The above diagram gives the impression that in order to recover their real and natural status in life, Hermione and Perdita had to undergo the formidable experience of reconciliation between their received/disguised existence and their uncovered/recovered real beings. The diagram might also express the fact that the displayed process necessarily opts for dynamism against stasis. We could also notice this dynamism in the Classical-to-Romantic shift. So, based on Joughin's introduction to *Philosophical Shakespeares*, "artistic creativity is crucially implicated in producing new forms of social interaction and in helping us to modify the criteria by which we understand their significance" (Joughin 4). Despite this, we do not suppose that Shakespeare intended to advocate the Romantic art form either. The first reason is the bridge he constructs between the Classical and the Romantic artistic forms discussed above. The second reason is related to his mocking treatment of the sublime and its Romantic manifestations.

If we propose that Shakespeare might have had human being in mind as a possible Romantic art form, we should add that he demystifies the form after all. The parodic treatment of human being as a Romantic art form is closely connected to Hermione. She is the only person literally undergoing a human-being-to-statue-to-human-being metamorphosis. In "Harry Hunks, Superstar", Terence Hawkes discusses the motif of bear-baiting (a pseudo-circus sort of performance popular at the time of Shakespeare and a rival to the theatre) in a number of Shakespeare's plays. With respect to *The Winter's Tale*, Hawkes attributes the motif to Hermione's condition. In its historical context, the motif meant a kind of inhumane and savage
play upon a tied and defenseless bear attacked by dogs or beaten by whips and so forth. According to Hawkes, Hermione suffers the same condition when she and Polixenes are talking silently while Leontes is nourishing extraordinarily superfluous jealousy in his mind against them, giving an account of "their behavior as startling in its immediacy as it is pointed in its implications" (Hawkes 96):

    Too hot, too hot!
    To mingle friendship far, is mingling bloods.
    I have tremor cordis on me,—my heart dances;
    But not for joy,—not joy…
    […]
    But to be paddling palms and pinching fingers,
    As now they are; and making practiced smiles,
    As in a looking-glass ….
    …O, that is entertainment
    My bosom likes not, nor my brows!–Mamillius,
    Art thou my boy? (I. II. 32-44)

The excerpt vividly evinces the effects of Leontes's jealousy giving "fractures" to his speech in declaring his wife's sexual betrayal (De Grazia 61).

Still, another evidence of bear-baiting motif happens in the third act, in the court of justice, where Leontes indicts the defenseless Hermione of lechery and desecrates her publically. Finally, the last scene where everybody is excitedly waiting for Hermione's statue to be uncovered, the motif is reiterated once more. Here, Hermione is standing on her pedestal, silent, motionless and fastened with onlookers pouring forth their comments on her aesthetic perfection. Of course, one might object to Hermione's being a statue in a context where we are talking about human being as a form. However, since in this scene Hermione serves both roles simultaneously (being a woman realistically and a statue virtually), the objection will be of no consequential significance.

What we get from the above evidences is a human being dragged into conscious
or unconscious, playful or earnest derision, confusingly idolized as a saint on the one hand, and desecrated and banished on the other. It seems that something is lacking in both contexts and that is realism. In other words, men and women in this play are, all the time, disguised in one or another way. As David Daniell also has marked, acting and role-playing could be seen everywhere in the play (Daniell 119) and, having Susan Synder's statement in mind, role-playing creates "a certain distance between audience and stage action which is increased by the conscious fictionality of that action" (Synder 95). So although human being is highly spiritual and sublime, s/he can easily take up different forms of disguise on various occasions. Disguise affirms the Lacanian notion that the realm of the real is beyond the reach of humanity. We are either still in the mirror stage or enter the symbolic order, but we can never attain the real.

The theme of disguise which derides the sublime has been exploited very brilliantly in the play. As a case in point, one can point to Autolycus' free slippage through the nets of caste and class benefiting from disguise of one sort or another. His convenient exchanging of clothes, now, as a vagabond, now, as a knight easily leads to an understanding of the potentially paradoxical nature of man (as the new romantic art form) who, endowed with sublimity, can embrace subversion according to the requirements of circumstances. Thus, cross-dressing allows Autolycus to violate the limits of class and get "something for nothing" (Hamilton 1993, par. 43). Autolycus has mastered the skill of make-believe which, in Bristol's words, "is characterized both by correct understanding of certain complex stipulations and by the way those stipulations are taken up in relation to the background knowledge of everyday life" (Bristol 27). Disguise has been conveyed in the following lines in a very sarcastic but significant way:

Unbuckle, unbuckle.—

[Florizel and Autolycus exchange garments.
Fortunate mistress,—let my prophecy
Come home to ye!—you must retire yourself

or unconscious, playful or earnest derision, confusingly idolized as a saint on the one hand, and desecrated and banished on the other. It seems that something is lacking in both contexts and that is realism. In other words, men and women in this play are, all the time, disguised in one or another way. As David Daniell also has marked, acting and role-playing could be seen everywhere in the play (Daniell 119) and, having Susan Synder's statement in mind, role-playing creates "a certain distance between audience and stage action which is increased by the conscious fictionality of that action" (Synder 95). So although human being is highly spiritual and sublime, s/he can easily take up different forms of disguise on various occasions. Disguise affirms the Lacanian notion that the realm of the real is beyond the reach of humanity. We are either still in the mirror stage or enter the symbolic order, but we can never attain the real.

The theme of disguise which derides the sublime has been exploited very brilliantly in the play. As a case in point, one can point to Autolycus' free slippage through the nets of caste and class benefiting from disguise of one sort or another. His convenient exchanging of clothes, now, as a vagabond, now, as a knight easily leads to an understanding of the potentially paradoxical nature of man (as the new romantic art form) who, endowed with sublimity, can embrace subversion according to the requirements of circumstances. Thus, cross-dressing allows Autolycus to violate the limits of class and get "something for nothing" (Hamilton 1993, par. 43). Autolycus has mastered the skill of make-believe which, in Bristol's words, "is characterized both by correct understanding of certain complex stipulations and by the way those stipulations are taken up in relation to the background knowledge of everyday life" (Bristol 27). Disguise has been conveyed in the following lines in a very sarcastic but significant way:

Unbuckle, unbuckle.—

[Florizel and Autolycus exchange garments.
Fortunate mistress,—let my prophecy
Come home to ye!—you must retire yourself
Into some covert: take your sweetheart's hat,
And pluck it o'er your brows; muzzle your face;
Dismantle you; and, as you can, disdain
The truth of your own seeming; that you may—
For I do fear eyes over—to shipboard
Get undescried. (IV. IV. 15-23)

A diverse type of subversion happens to Perdita who, disguised under a coercively rustic façade of peasantry for sixteen years, casts off her peasant clothes and wears those of a princess, joins the royal family through the adoption of a new identity, new parents and a new life style. Perdita's disguise is clearly discernible to Polixenes when he meets her for the first time in the Shepherd's house; he easily reads nobility in her features and gets full of admiration for her:

This is the prettiest low-born lass that ever
Ran on the green-sward: nothing she does or seems
But smacks of something greater than herself,
Too noble for this place. (IV. IV. 180-83)

Moreover, Paulina and Hermione play make-believe roles as conjuror and sculpture quite cunningly. They are totally aware of their companions' emotional status having a sixteen-year store of background knowledge before they impart the shock of revealing the truth. They manipulate their feelings dexterously. The following excerpt shows Paulina's make-believe role by announcing herself capable of doing magic things and receiving help from wicked powers:

Either forebear,
Quite presently the chapel, or resolve you
For more amazement. If you can behold it,
I'll make the statute move indeed, descend
And take you by the hand: but then you'll think,—
Which I protest against,—I am assisted
At this juncture, regarding make-believe or disguise, we should make a distinction among Autolycus, Paulina, Hermione and Perdita. The first deploys the above sleights of hand quite consciously and even candidly acknowledges what he has done to the audience. This consciousness exists in Paulina's case too.

Hermione's condition, on the other hand, is a bit formidable. Her adoption of disguise can be interpreted differently under diverse circumstances; as to her vanishing for sixteen years, we should mark that she is coerced to non-existence by the King, so though conscious of her situation, she plays this make-believe role quite reluctantly at the initial stage. As for her conversion into a statue, we feel more confident to say that she plays the role very willingly if we consider her longing to return to her former status. While reading the play, we feel as if Hermione had been defamed and winning back her honor would not be a simple task; if we take Umberto Eco's outlook on Greek or Classical rationalism into account, truth is something both explicable and inexplicable. The inexplicability of the knowledge of truth is due to the fact that it is always beyond the reach of the Classical man and in gods' possession; attaining truth, therefore, presupposes attaining gods' language (i.e. a different language) (Eco 30). That is why Leontes waits for Apollo's oracle, it is the ultimate and indubitably true judgment, since, first of all, its secrecy and then its revelatory certainty is beyond man's power. But Leontes doubts the oracle's authority and his son is taken away from him by gods as a punishment (one of the reasons for maintaining that the play shifts from the moderation of Classicism to the insurrection of Romanticism is that the boundaries are broken and revolt dominates the atmosphere, in this case, Leontes's revolt against gods). So, the time is not ripe yet for Hermione's restoration of her previous glory. This needs a much braver attempt on her part. To reveal the truth of her honor, Hermione must adopt a different language, the language of aesthetics, namely sculpture.

Perdita's case is a horse of a different color. She was disowned by her father, found in the wilderness by chance and raised by a shepherd; she fell in love with a
prince unaware of the fact that his father was a friend of Leontes (her real father); but could Polixenes also be unaware of the fact that he was actually conversing with Leontes's daughter despite her striking likeness to her mother? The point is that Polixenes's negligence has been exposed by Shakespeare implicitly, since he does not recognize "her strong resemblance to Hermione, which is stressed as one of the 'proofs' of identity later" (Thatcher par. 4) on in the play. This could be an evidence of the insufficiency of the Classical notion of exact imitation (e.g. the striking resemblance between Perdita as daughter and Hermione as mother) which has been included and ridiculed in the play quite indirectly by the author. One more evidence of Polixenes's conspicuous negligence is that the fact that Perdita cultivates flowers in her garden "hints at her desire for upward mobility or, perhaps more likely, at her hidden royal heritage." This, however, does not lead to Polixenes's recognition of Perdita (Tinger 126). by means of a very hasty elopement with Florizel, she accidentally finds herself in her father's abode. As a result, she is not conscious of the imposed role she has been playing all these years.

The above details may prove that Shakespeare was too smart to fall prey to any of the art forms classified by Hegel. We might even say that by commingling both forms at first, and then parodying them, Shakespeare caused violation in the linear, progressive movement from Classical to Romantic aesthetic notions. As it is, therefore, *The Winter's Tale* may be understood to contain anti-Hegelian perspectives.

**Conclusion**

To conclude, we should point out that the amalgamation and hybridization of the two themes of "disguise" and "reconciliation between Classic and Romantic aesthetic forms" in this play might be a good reason for the recognition of Shakespeare's attitude toward those aesthetic notions as paradoxical and ambiguous. Thus, back to the questions we raised earlier, and with respect to Shakespeare's time as the age of Reformation, we could state that *The Winter's Tale* seems to display a movement from Classical to Romantic aesthetics at the first glance. But then,
delving into the heart of the matter, we find out that the play deals with both aesthetic forms so ambivalently that their mutual embracement and reciprocal interaction wear away to the point of uncertainty. We can definitely detect both Hermonian stasis and resurrection/insurrection in the play, but we should be keen enough to notice the turbulence beneath those notions. As for the stance of this essay, we may agree with Joughin about the paradoxical nature of Shakespeare's works resisting any sort of generalization. As a result, general statements such as "Shakespeare is the classic of motion" should be approached more cautiously.

References


Gender Differential Item Functioning Analysis of the University of Tehran English Proficiency Test

Abbas Ali Rezaee*  
Assistant professor, Faculty of Foreign Languages and Literatures, University of Tehran, I.R. Iran

Enayatollah Shabani**  
PhD student of TEFL, Faculty of Foreign Languages and Literatures, University of Tehran, I.R. Iran  
(Received: 8 Dec. 2009, Accepted: 4 Feb. 2010)

Abstract
The University of Tehran English Proficiency Test (UTEPT) is a high-stakes entrance examination taken by more than 10,000 master’s degree holders annually. The examinees’ scores have a significant influence on the final decisions concerning admission to the University of Tehran Ph.D. programs. As a test validation investigation, the present study, which is a bias detection research in nature, utilized multistep logistic regression (LR) procedure to examine the presence of gender differential item functioning (DIF) in the UTEPT with a sample of 6,555 examinees who took the test in November 2006. Specifically, the LR DIF two-degree of freedom Chi-squared test of significance was employed to test the significance of DIF. Following what has been recently suggested in the literature, the test of significance for DIF was accompanied by a measure of magnitude, namely the $R^2$-squared effect size for LR DIF, for the interpretation of which the two widely accepted classification schemes were used comparatively. The results reveal that 39 of the 100 items in the test display significant gender differences. However, these group differences are viewed as “negligible” based on both of the schemes. Accordingly, it could be argued that the UTEPT is a gender DIF-free test, though the Reading Comprehension section of the test remains in need of further analysis as it seems that the general trend of DIF indices at the item level may hint at an inclination towards males.

Keywords: Differential Item Functioning, Logistic Regression, UTEPT, English Proficiency Test, Gender Differences, High-stakes Test.

* Tel: 021-61119081, Fax: 021-88634500, E-mail: aarezee@ut.ac.ir  
** Tel: 021-61119081, Fax: 021-88634500, E-mail: shabani@ut.ac.ir
I. Introduction

One of the important factors which should be taken into account in dealing with the validity of any test is the issue of fairness (Thissen, 2001). Educational Testing Service Fairness Review Guidelines (2003) offers a simple and fairly straightforward verdict: “A test that shows valid differences is fair; a test that shows invalid differences is not fair” (p. 2). Therefore, when people with similar abilities in the construct being measured perform substantially differently on a test item it is necessary that the item be reviewed for fairness (Gierl, Khaliq & Boughton, 1999), and perhaps removed if the differential performance of the examinees is not balanced or cancelled over the test as a whole (Wainer, Sireci & Thissen, 1991). This issue has made a great area of investigation for researchers, and a technique which empirically measures the differential functioning of items for groups which are matched on the ability of interest has now become “the new standard in psychometric bias analysis” (Zumbo, 1999, p. 6) and is now “a key component of validity studies in virtually all large-scale assessments” (Penfield & Camilli, 2007, p. 125). This technique is called differential item functioning (DIF) after Holland and Thayer who used the term in a seminal chapter on test validity (1988).

DIF occurs when the responses of individuals having the same ability of interest show systematic differences simply based on their membership in a certain group. It is necessary, therefore, that judgmental and statistical analyses be applied to detect the items in a test that function differentially for individuals with the same ability from different groups. It can be argued that the higher the stakes of a test, the more the need for such analyses. The necessity of conducting a study examining the issue of fairness in the University of Tehran English Proficiency Test (UTEPT) is best understood when one takes into account the fact that every year the test is taken by more than 10,000 masters degree holders and the inferences made on the basis of the test scores are a crucial factor in determining the admission of the test takers in the Ph.D. program at the University of Tehran.

Gender DIF analysis is an important part of test development as it helps to examine and eliminate the items which may be potentially unfair to some groups of
Gender Differential Item Functioning ... 91
test takers because of gender group membership. To the knowledge of the present researchers, no such study has ever been done in Iran. The purpose of the present study, therefore, is to investigate the presence, the magnitude as well as the direction of DIF in the items of the UTEPT.

II. Background

DIF exists when after controlling for overall ability examinees from different groups have a different probability or likelihood of successfully answering an item. Though the development of statistical methods for identifying potentially biased items began in the 1970s (Penfield & Camilli, 2007), it seems that the first bias studies were done due to the Civil Rights Act of 1964 in the U.S. (Conoley, 2003 and Duncan, 2006). Owing to the Act, the first DIF studies were those which were conducted to ensure test fairness for black and white examinees. The methods used in the early DIF studies were ANOVA, correlational methods and Transformed Item Difficulty (TID) method which looked for differential difficulty as an indicator of bias. Since that time, many approaches have been developed to examine DIF, among which standardization (Dorans & Kulick, 1986), Mantel-Haenszel (MH; Holland & Thayer, 1988), logistic regression (LR; Swaminathan & Rogers, 1990), Simultaneous Item Bias Test (SIBTEST; Shealy & Stout, 1993) and item response theory (IRT) procedures are the most widely used statistical methods for detecting DIF in dichotomously scored responses (see Sireci & Allalouf, 2003 for a detailed list of DIF methods and the types of data for which each method is appropriate).

Generally, DIF is of two types, uniform and non-uniform. Uniform or unidirectional DIF exists when the probability of endorsing an item (answering an item correctly) is greater for one group than for the other group over all the levels of proficiency. The existence of non-uniform or crossing DIF demonstrates that the difference in probabilities of a correct response is not the same at all levels of proficiency between the two comparison groups. That is, the probability of correctly answering an item is higher for one group at some points on the scale, and higher for the other group at other points (see Salkind, 2007). In IRT terminology, non-uniform
DIF exists when the item response curve for the focal group intersects the item response curve for the reference group. If the item response curves for the focal and reference groups do not cross, it is said that DIF is uniform. In other words, uniform DIF occurs when there is no interaction between the ability level and group membership. If there is an interaction between ability and gender, this signals the existence of non-uniform DIF (Swaminathan & Rogers, 1990).

For many years, the prevalence of non-uniform DIF was an intense debate among researchers. This explains the introduction of a DIF examination procedure for both uniform and non-uniform DIF as late as 1990, that is, almost three decades after the first statistical methods for DIF detection were proposed. Though the controversy seems not to have ended up in a unanimous agreement, there is little doubt today as to the occurrence of non-uniform DIF (Jodoin & Gierl, 2001). Therefore, utilizing an approach which can detect both uniform and non-uniform DIF seems superior for the researchers to enjoy a higher standard of practice. With this aim in view, the LR procedure was used as the DIF detection procedure in the present study.

III. Previous Gender DIF Studies in Language/Verbal Tests

Though the issue of bias has long been a concern of language testers, the systematic study of it can only be observed in the studies conducted after 1960. Yet, it was only as late as the 1980s that with the advent of more sophisticated approaches to DIF examination researchers could screen language tests for what could possibly lead to the unfair treatment of examinees of different gender. In one of these studies, Lawrence and Curley (1989) found that the items related to technical science content were more difficult for females. Curley and Schmitt (1993) found that of the 2,250 SAT-Verbal questions pretested in 1990, only 8.5 percent displayed moderate and large amounts of gender DIF. Buck, Kostin and Morgan (2002) found that item content can be associated with gender-based performance differences. In a study conducted by Ryan and Bachman (1992), it was found that males and females do not react differently at the item level in TOEFL and FCE. Lin
and Wu (2003) conducted a DIF study on the EPT (English Proficiency Test) in China, which was modeled after the TOEFL and found that there was not much gender DIF in the test. As they reported, of a total of 120 items, only two items were identified as exhibiting C-level or large DIF and in fact 89% of the test items exhibited “no” DIF whatever.

Conoley (2003) conducted a study to detect DIF in a measure of receptive vocabulary and found that expert judges could not accurately predict items that demonstrate DIF against one group, and that it was more plausible not to rely heavily on the expert’s subjective speculations in bias detection studies. The important point to note here is that DIF is a test-dependent phenomenon. Differential performance of the members of different gender groups on a certain item in a certain test does not reveal anything about the existence or direction of DIF in other language tests. In order to examine DIF in a test, an objective DIF analysis must be done on exactly that very test, and the findings of other DIF studies do not indicate much about the presence or absence of DIF in the test. Taking this point into account, the present research was an attempt to investigate the presence, magnitude, and direction of gender DIF in the UTEPT.

IV. Method

1. Data Set

The data used in this study consisted of the test results from the UTEPT which was taken by master’s degree holders of different majors seeking entry to the University of Tehran Ph.D. program in the year 2006. The analysis was performed on 6,555 cases, 4553 (69.5%) males and 2002 (30.5%) females. The male group was considered as the reference group and coded 1, and the female group was taken as the focal group of the study, given the dummy code 2. (Though McNamara and Roever, 2006, consider such categorizations conceptually problematic, it is common practice in DIF studies.)

Broer, Lee, Rizavi and Powers, (2005) presume a large sample size to be a
requirement for the LR procedure. Schmitt, Holland and Dorans (1992) also suggest that the largest possible number of examinees should be used to obtain stable DIF estimates and ensure sufficient power for DIF detection. Curley and Schmitt (1993) recommend that in order to come up with reliable results, large sample sizes be used, particularly for the focal group. Therefore, the sample size of the present study seems to enjoy a considerable degree of acceptability for the results to be reliably interpretable.

2. UTEPT

The University of Tehran English Proficiency Test (UTEPT) is a test of English proficiency developed in the Language Testing Center of the Faculty of Foreign Languages and Literature, University of Tehran. It functions as a screen test which must be taken by the M.A./M.S. holders of different majors to be allowed to sit the major-specific Ph.D. entrance examination. The test is taken by more than 10,000 master’s holders annually and the final decisions concerning admission to the University of Tehran Ph.D. candidacy are influenced by the results of this test. The test calibrates items separately within subtests and creates its final scale as a weighted composite of the scales created within subtests. However, the scores for the separate subtests are available and can be used for research purposes.

The UTEPT is a set of 100 four-option multiple-choice items with a time limit of 100 minutes. It includes three parts: The Structure and Written Expression section which comprises 40 items, the Vocabulary section which consists of 25 items, and the Reading Comprehension section which has 35 items. Throughout the study, care was taken to consider the developers’ concern over the confidentiality of the test.

The development process of the UTEPT includes three main steps. In the first step, the items are written. Then their usefulness is evaluated through different stages of review, and finally the final form of the test is created. There is yet another stage of internal review which finalizes the selection of test items and is applied after the test items are selected. Although biased items might be identified through the review stages of test development, statistical procedures can provide the test
developers and users with a robust niche objectively and empirically as regards the fairness of the test.

3. Procedure

An approach first proposed by Swaminathan and Rogers (1990), the LR approach to examining DIF in dichotomous items is based on the examination of regression in predicting item response from main effects of group and ability and their interaction (Kanjee, 2007). In other words, in binary LR, the item response, which in binary scoring format can have two values (correct or incorrect), will be the dependent variable, and the grouping variable (e.g., male vs. female), the ability variable and a group-by-ability interaction term will be taken as independent variables.

In the present study, the dichotomously scored item-level data based on 6,555 examinees was subjected to the multistep LR model which uses group membership \((g)\) and ability-by-interaction \((g*\theta)\) term to estimate the probability that a randomly selected examinee with an ability of \(\theta\) answers an item correctly. The LR equation can be shown as
\[
Y = b_0 + b_1\theta + b_2g + b_3\theta g,
\]
where \(\theta\) is the ability level of the examinee (which is usually represented by the total scale score), the letter \(g\) denotes the group membership variable which can take the two codes specified for the reference and focal groups, \(b_0\) is the intercept for a dichotomized variable, and the parameters \(b_1\), \(b_2\) and \(b_3\) are the slope parameters which denote the weights for ability, group membership, and the interaction of the two, respectively (Zumbo, 1999). The asterisk represents interaction between variables. In fact, this equation is a linear equivalent of the original LR DIF formula proposed by Swaminathan and Rogers (1990). The null hypothesis in LR DIF analysis can be shown as
\[
H_0: b_2 = b_3 = 0
\]
for each item.

The LR method for binary scores employed in the present study was the three-step modeling process proposed by Zumbo (1999). In this method, in the first step the conditioning variable \((\theta)\) is entered into the regression equation. Conditioning is one of the requirements for every comparison between the groups. In the second
step, the group membership variable (g) is entered into the equation, and finally, in the last step, the score-by-group interaction term (\(\theta \times g\)) is entered. In sum, the whole process can be described as the following:

Step 1: enter the conditioning variable (\(\theta\))

Step 2: enter the group membership variable (g)

Step 3: enter the term for the interaction between the conditioning variable and group variable (\(\theta \times g\))

In practical terms, by subtracting the Chi-squared value of the model with the ability variable from the Chi-squared value of the model with the interaction term and comparing the results with its distribution function with 2 degrees of freedom, the statistical tests for DIF can be computed (the two degrees of freedom is the result of comparing the model Chi-squared statistic at Step 3, which is three, and the model Chi-squared statistic at Step 1, which is one). An item is said to exhibit significant group differences when the \(p\)-value for the Chi-squared test of significance for that item is less than or equal to .01.

However, simulation studies have indicated that using the LR procedure without a measure of effect size could result in situations in which Type I error rate would be higher than expected (Jodoin & Gierl, 2001). In fact, an effect size measure is a descriptive statistic representing the “magnitude” of DIF. Since different studies have suggested that the power of statistical test is sensitive to the employed sample size (e.g. Jodoin & Gierl, 2001 and Zumbo, 1999), it seems essential that a measure of magnitude be used to enable the researcher to interpret DIF in terms of size.

The effect size for LR DIF is tested by the \(R^2\)-squared coefficient. \(R\) is a statistic that is used to look at the partial correlation between the dependent variable and each of the independent variables. \(R\) can range from -1 to +1. Within this framework, \(R^2\) changes are used to represent the magnitude of DIF. \(R^2\) is the result of the ratio of model fit Chi-squared to -2Log likelihood. In effect, the \(R^2\) values corresponding to uniform and non-uniform DIF are associated with \(b_2g\) and \(b_3\theta \times g\), respectively. In this procedure, uniform DIF can be measured by comparing the \(R^2\) value of the ability-only regression equation (i.e., Step 1) with the \(R^2\).
value after adding the group variable (i.e., Step 2). Non-uniform effect sizes can be calculated by comparing the value of $R^2$ in Step 3 and Step 2, that is, after and before entering the interaction term into the LR equation. It is the aggregation of uniform and non-uniform effects that makes the total DIF effect size.

Zumbo and Thomas (1996) offer a classification guideline for DIF effect sizes. Based on this guideline, DIF values can be classified as negligible (A-level), moderate (B-level) or large (C-level). This guideline can be summarized as the following:

1. Negligible or A-level DIF: $R^2 < 0.13$
2. Moderate or B-level DIF: $0.13 \leq R^2 < 0.26$
3. Large or C-level DIF: $0.26 \leq R^2$

Therefore, putting the two procedures together, as Zumbo (1999) points out, for an item to be classified as displaying DIF, the two-degree-of-freedom Chi-squared test in logistic regression had to have had a $p$-value less than or equal to 0.01 (set at this level because of the multiple hypotheses tested) and the Zumbo-Thomas effect size measure had to be at least an $R$-squared of 0.130. (p. 27) [italics original]

It should be noted, however, that this is not the only classification scheme for $R^2$ values in the DIF literature. In a simulation study on evaluating Type I error and power rates using an effect size with logistic regression DIF, Jodoin and Gierl (2001) compared the effect size measures for logistic regression and SIBTEST to suggest a different set of classification standards for $R^2$ changes, which they believed could be considered strikingly different form the one proposed by Zumbo and Thomas (1996). In brief, the guideline proposed by Jodoin and Gierl can be summarized as the following:

1. Negligible or A-level DIF: $R^2 < 0.035$
2. Moderate or B-level DIF: $0.035 \leq R^2 < 0.070$
3. Large or C-level DIF: $0.070 \leq R^2$

We used both of these classification schemes in this study.
4. Analysis

In this study the multistep LR procedure was used to examine the existence, the magnitude, and the direction of DIF in the items of the UTEPT. For this purpose, first the grouping variable needed to be coded and defined as nominal. Dummy codes 1 and 2 were used for the reference and focal groups respectively. Also, the total score for each individual was needed to be specified to be used as the conditioning variable. The item responses were then coded dichotomously. That is, all the 100 item responses for the entire 6,555 examinees were converted to a 0-and-1 system of coding.

The dependent variable (item response) and the independent variables (ability and gender) were entered into SPSS and the binary logistic analysis was used to analyze the data. For each item, the simultaneous uniform and non-uniform DIF two-degree of freedom Chi-squared (DIF $\chi^2[2]$) was produced, which was the test of significance for DIF in LR. The $R^2$ changes for the items with significant group differences were then estimated and compared against the Zumbo-Thomas (1996) and Jodoin-Gierl (2001) classification schemes to find the items with moderate or large DIF indices.

V. Results

Of the 6,555 test takers whose records have been used in the present study, 4553 were male and 2002 were female. The average total score of all the participants in the test was 54.71. The means of the total scores for males and females were found to be 54.12 and 56.04 respectively. A preliminary analysis of the data indicated that the standardized mean difference in total scores of males ($M = -.036$, $SD = 1.01$) and females ($M = .083$, $SD = .96$) was $d = -.12$, which was found to be statistically significant, $t(6553) = -4.47$, $p \leq 0.01$ (two-tailed). However, it would be viewed as a very small effect size using Cohen’s (1988) standard of .20 as “small”. Table 1 below gives a summary of the descriptive statistics.
The results of the formal test of significance for DIF revealed that of the 100 items in the test, 39 items could pass the \( p \leq 0.01 \) condition for the two-degree-of-freedom Chi-squared test. Table 2 presents results of the DIF analyses for these 39 items. The results indicated that non-uniform DIF was not prevalent in the UTEPT and from the 39 items which could pass the test of significance, 24 items were found not to exhibit any non-uniform DIF whatever. This means that 85 percent of the items in the entire test displayed no non-uniform DIF. On the whole, items 51 and 83 exhibited the largest gender effect sizes, \( R^2=0.01 \) (see Table 2).

### Table 2. Uniform, Non-uniform, Total R² Effect Sizes, and the Chi-squared test results for the items with \( p \leq 0.01 \)

<table>
<thead>
<tr>
<th>Item No.</th>
<th>Uniform ( R^2 )</th>
<th>Non-uniform ( R^2 )</th>
<th>Total ( R^2 )</th>
<th>( \chi^2 ) (2)</th>
<th>( p )</th>
</tr>
</thead>
</table>

**Table 1. Descriptive statistics of the male and female test-takers’ scores**

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Mean of Total Score</th>
<th>Standard Mean Difference</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>4553</td>
<td>54.12</td>
<td>-.036</td>
<td>1.01</td>
</tr>
<tr>
<td>Female</td>
<td>2002</td>
<td>56.04</td>
<td>.083</td>
<td>.96</td>
</tr>
</tbody>
</table>
VI. Discussion and Conclusions

The present study was an attempt to examine the fairness of the UTEPT with regard to gender. It can offer some insights into the test design and development process at the University of Tehran. Although there are some researches to suggest that multiple choice questions might favor either boys or girls (Woods, 1991), the general consensus is that at the item level, no group differences should be found in a test of English language proficiency. Gender group membership should not have any impact on the performance of examinees on an item. This is of great prominence
since the UTEPT is a high-stakes test with a major impact on people’s future academic and social life. Though the existence of DIF in a test does not necessarily signal bias, the present study was conducted as a prelude to more profound analyses towards developing unbiased tests.

The results revealed that of the whole 100 items in the test, 61 items could not pass the formal test of significance for DIF. That is, the two-degree of freedom Chi-squared test for these items did not have a \( p \)-value less than or equal to .01. This left 39 items (39% of the test items) to be tested for the magnitude of DIF. Once the effect sizes for the items with significant gender differences were estimated, these values needed to be interpreted based on the classification schemes for DIF. If an item with significant group differences exhibited \textit{moderate} or \textit{large} DIF, it could be considered as a potential threat to the fairness of the test. In other words, it can be suggested that these test items should be replaced with ones which show less DIF.

Based on both Zumbo-Thomas (1996) and Jodoin-Gierl (2001) guidelines, all of the 39 items with significant gender differences were found to display negligible or A-level DIF. In consequence, it can be concluded that concerning gender the UTEPT appears to be a DIF-free test. This, however, can only be regarded as a general conclusion since it only takes the aggregate results into account. It is not implausible to argue that 39 out of 100 is still a lot, though the magnitudes for DIF indices were shown not to have reached the acceptable level to be considered moderate or large. What is of equal importance in interpreting the results is the individual direction of DIF indices and the cumulative effect of these indices. It seems necessary therefore to examine whether these “negligible” DIF values display random preferential treatment for either of the gender groups or it is the case that a considerable number of the indices favor one group more than the other, in which case, the argument supporting that the test is gender DIF-free can hardly remain persuasive.

The individual analyses of the 39 items which could pass the test of significance for DIF displayed that 25 items favor males and 14 items favor females. This suggests that on the whole these “negligible” DIF indices are shown to have favored
males more than females. It is even possible to have a closer look at the direction of the DIF values by considering the three parts of the test and comparing the direction of the items that favor males and those that favor females in the three sections, namely Structure and Written Expression, Vocabulary, and Reading Comprehension.

Out of 40 items in the Structure and Written Expression Structure section, 11 items contained DIF values which could pass the formal test of significance. From these 11 items, six favored males and five favored females. Taking the Vocabulary section into account, it was found that out of the 25 items in this section, 13 items displayed significant DIF values, eight favoring males, and five favoring females. Finally, in the Reading comprehension section, 15 out of 35 items in the section were found to have significant DIF indices, out of which 11 favored males and four favored females.

Putting the magnitude of the found DIF values aside for the moment, this subsequent exploratory analysis of the results reveals that with regard to the differential performance of males and females on different sections of the test, it is the Reading Comprehension section of the test which for the most part appears to be responsible for the differential performance of males and females on the items of the UTEPT (note again that the group differences were found to be significant, but negligible). It might be useful then to have a closer look at this section of the exam.

The Reading Comprehension section of the test comprises various questions which can generally be divided into three parts. The first part, which is the main part of the section, contains 28 questions and consists of six passages, ranging from 87 to 233 words, each followed by four to six multiple-choice questions. The second part which includes only four questions is called the Restatement part in which the test takers were supposed to read sentences followed by four choices and choose the best restatement of the sentences. The last part is called Coherence and has just three questions. For each item in this part, the test takers were required to read a paragraph in which one sentence had been removed and choose the choice that best completed the paragraph and made it coherent.
A cursory examination of the items with significant DIF in the Reading Comprehension section revealed that passages 1, 4, 5 and the Coherence part were mostly in favor of males. Though the DIF values were found to be “negligible”, the finding that all of the significant DIF values related to the items following these three passages and the Coherent part favored males should not be taken as accidental. In general, the themes of the three passages favoring males can be summarized as “educational development”, “education of children with disabilities”, and “the International Bank”, respectively. Nonetheless, it should be admitted that drawing any conclusions at this point is neither helpful nor reliable. In fact, a more careful and detailed analysis of the reading passages is needed before getting to any generalizations.

By and large, this study posed three research questions to investigate: First the presence, second the magnitude, and third the direction of any gender DIF in the items of the UTEPT. With respect to the first and second research questions, it was found that of the 100 items in the test, 39 items had a significant $p$-value for their test of significance; however, the detected DIF indices are considered “negligible” using the corresponding classification schemes for DIF effect size proposed by Zumbo and Thomas (1996) and Jodoin and Gierl (2001).

The direction of DIF can be estimated from the regression coefficients, or particularly from the gender group main effect $b$-weight or the slope parameter of the regression. In fact, the direction of DIF depends on the sign (positive or negative) of the group $b$-weight and how group membership is coded (B. Zumbo, personal communication, August 1, 2008). Although some efforts were made in the present study to explain the small differences between the performances of the two genders on some items (particularly on the Reading Comprehension section), it should be noted that the direction of DIF is interpretable only when there is a DIF at all. Therefore, since no DIF index at the item level was detected in the test, this question needs to be examined in a more comprehensive study which takes differential test functioning (see Pae & Park, 2006) or differential bundle (testlet) functioning into consideration (see Wainer et al. 1991).
The results of the study can bear more significance by taking one point into account. LR is a parametric DIF detection approach which is a response to the previous DIF techniques which could only screen uniform DIF such as standardization, MH or SIBTEST. In a study comparing LR and MH for detecting DIF, Rogers and Swaminathan (1993) concluded that the LR procedure was as powerful as the MH procedure in detecting uniform DIF, and more powerful than the MH in detecting non-uniform DIF (see also Gierl, Jodoin & Ackerman, 2000). In addition, as Duncan (2006) stated,

If LR DIF can detect non-uniform DIF better than the MH DIF method, and is as powerful at detecting uniform DIF as the MH DIF method, then the inclusion of an effect size would make LR DIF a very attractive choice as a DIF detection method. (p. 38)

Jodoin and Gierl (2001) also emphasized that while LR has comparable power to MH and SIBTEST in detecting uniform DIF, it is superior in power for detecting non-uniform DIF. In addition, Gierl, Rogers and Klinger (1999) found that “effect size measures [for MH, SIBTEST and LR] were highly correlated across DIF procedures except the measure for non-uniform DIF” (p. 15), which could only be assessed by LR. Altogether, these findings can provide tacit confirmation as to the superiority of LR over standardization, MH and SIBTEST.

There is another point which may add to the value of the findings of the study. In an attempt to examine the reliability of different DIF detection methods, Huang (1998) made a comparison between standardization, MH and LR methods and found that more items could be identified as exhibiting DIF by LR than the MH and standardization methods. That is, “items which were labeled as ‘exhibiting DIF’ by the MH and STD [i.e., standardization] methods could be identified as either uniform DIF or non-uniform DIF in the LR method” (p. 8). In other words, all the items labeled as “exhibiting DIF” by both the MH and standardization methods, were also detected to exhibit DIF by the LR method. It can be concluded, therefore, that in detecting DIF items “the LR method is more sensitive than the MH and STD methods” (p. 8), and that in comparison with MH and standardization, the LR
method tends to label the most items as exhibiting DIF. This is consistent with the findings of Gierl, Khaliq, et al. (1999), who found that in comparison with MH, which they referred to as “the conservative procedure” (p. 12), LR could flag a large number of items as exhibiting DIF. Also, as Gierl, Jodoin, et al. (2000) have found, LR has excellent Type I error rates which is a reassuring point for the researchers who choose LR as their DIF detection method.

On the whole, what can be inferred from this comparative discussion is that, generally, LR is more likely to flag an item with moderate or large DIF than the other two DIF detection methods which have been generally used at ETS viz standardization and MH. In other words, by utilizing LR, the researchers can be sure that they obtain a list of DIF items which might not be flagged as displaying DIF by either the MH or standardization procedures. Metaphorically, LR feels free to accuse an item of displaying DIF (see Duncan, 2006; Gierl, Jodoin, et al., 2000; Gierl, Rogers, et al., 1999; Jodoin & Gierl, 2001 and Rogers & Swaminathan 1993).

This, implicitly, can be considered as a reassuring point for the developers of the UTEPT. In other words, since the procedure which was used in the study is considered to be the strictest in comparison to standardization and MH, and since LR could not flag any item as displaying DIF in the test, the developers of the UTEPT can be convinced that it is most probable that the items of the test would not be flagged as displaying DIF if another study were to be conducted using standardization or MH. In this regard, a better plan for future studies may be to conduct another DIF analysis on the test utilizing item response procedures which seem to be the most efficient methods of all. For one thing, unlike MH and standardization, IRT is capable of screening non-uniform DIF which is important if the follow-up studies are to be comparative (see Kang & Cohen, 2007 for IRT model selection methods for dichotomous items). Generally, an ideal DIF study is the one which uses a combination of DIF techniques. Also, further research is needed to examine the differential functioning not of items but of the whole test or items bundles particularly the Reading Comprehension section. In addition, other studies can examine the existence of DIF for examinees with different academic
backgrounds or disciplines.

References


Islamic, Universal Concepts in the Shahnamah

Eshagh Toghyani*
Associate Professor, University of Isfahan, I.R. Iran
(Received: 8 July. 2007, Accepted: 19 Jan. 2010)

Abstract
As an Iranian literary work, it might seem that The Shahnamah is a national Epic which has been composed to promote Persian language and protect historical and cultural identity of Iranians against the dominance of alien cultures. However, this paper attempt to argue that the investigation of the stories, characters, and the symbols presented in the Shahnamah indicate that Ferdowsi's thought and ideas go beyond the borders of a single culture and cover a broader scope which encompasses the whole universe and human beings. This is related to the fact that the Shahnamah is deeply rooted in Islamic ideas and world view. In this view, we find that the Shahnamah is a splendid literary work which by presenting typical ideas, images and profound human experiences such as reason, love, justice, creation, life, death and invalidity of secular life, among many others, conveys universal Islamic messages rather than mere nationalism and racism. In fact, while Ferdowsi is exploiting ancient Iranian stories, the very mythical nature of these stories gives an opportunity to him to address all human beings. These features in the Shahnamah have ever since been the source of inspiration for different Iranian and non-Iranian artists to create valuable works in calligraphy, painting, poetry, story and play writing, music, cinema and other arts.

Key Words: Ferdowsi, Shahnamah, Epic, Cultural Identity, Myth, Universal Ideas.

* Tel: 0311-7933163, Fax: 0311-7933163, E-mail: etoghyani@hotmail.com
I. Introduction

Ferdowsi (932-1016 A.D.) is one of the greatest poets of Iran whose masterpiece, *Shahnamah*, brought him everlasting fame and glory. Since its composition (401 A.H.), many critics and scholars, in a wide range of books and articles, have investigated this work and have suggested different opposing views about Ferdowsi's life, thoughts, art and religion.

In most of these works it is suggested that Ferdowsi has created this work in order to save the national and cultural identity of Iranians and commemorate the race of Arya (Nodushan 1363 A.H.: 2). In this vein, they have tried to show that Ferdowsi's frame of reference has been nothing but nationalism, racism, and the religions of ancient Iran. The arguments usually center around the point that since most of the stories of the *Shahnamah* are about Iran and Iranians, thus, Ferdowsi's outlook is based on Zoroastriansim and/or Mazdaism (Meskoob 1356 A.H.: 69).

In this paper, however, by presenting several pieces of evidence, I will argue that, if we keep away from prejudice, we find that Ferdowsi is a writer who has portrayed more than other scholars. Rather, I will suggest that Ferdowsi's thought and art bear universally humanistic significance which directly originates from Islamic teachings.

2. Epic Poetry

The first point to mention is the fact that Ferdowsi addresses his audience using the lofty language of an epic. Considering the general properties of this genre and the wonders of its universe will lead us to a better understanding of Ferdowsi and his masterpiece.

"Epic being a long narrative poem is more purely a form of speech, related in an elevated style, whose subject matter consists of the exploits of heroic figures, men of great stature, whose exploits are given vast significance, affecting the fate of a tribe, nation, or the human race " (Mokhtari 1368 A.H.: 21). The hero is a figure of great national or even cosmic importance. The setting of the poem is ample in scale and may be worldwide and even larger (Abrams 1971: 51). The stories are derived from
mythical and historical materials developed in the oral tradition of a nation during a period of expansion and warfare.

*Shahnamah* as an example of this genre, can be divided into three parts: the mythical part, the heroic, and finally the historical. In the first part, the stories center around the creation of man and the course of his civilization. In this period the battle between evils and humans is the essence and main substance of the stories. Jamshid is the great king, who after defeating the evils, establishes human civilization. The discovery of fire, building the first human dwellings and medical sciences start in his reign. Hence, anarchy is cleared from the earth, and the humans forget the evils. This is the state of affairs until Jamshid becomes arrogant and conceited as a result of which Zahhak overthrows his monarchy and everything over turns.

His name Zahhak, and most ambitious he,
Light-happened, brave and fearless as could be…
Now it so happened Iblis on a day,
Com him to visit in a friendly way,
The chief’s heart frome the right path led astray,
The youth his ears bent to his words that day.
Of all his evil doing unaware,
He was well pleased his sweet discovere to share.
Conscience and heart to yield him he was led,
And thus was dust poured on the youth’s own head
(Ferdowsi, 1973: 22)

In the second part, which is the heroic era, the most important part of *Shahnamah*, great heroes arise. The greatest of these is Rostam, an ideal whose most perfect moral virtues have made him a perfect human being. Rostam is so unique that no character has ever resembled him in the literature of any language. He has been developed in the imagination of thousands of people in the course of time. A hero who has been created as the embodiment of their long standing dreams, wishes and ideals. The most important and magnificent heroic stories which appear in this
part of the *Shahnamah* were epic in its precise meaning, revealing its real beauty at its best. Everything such as time, place, episodes, and the like is extraordinary and supernatural, and the episodic nature of the plot, the lofty style and ornamental language, noble characters and supernatural events all contribute to this achievement.

Finally, in the third part, most stories convey historical events while the tone becomes less heroic.

### 3. Myth and Symbolism in the *Shahnamah*

As regards the first part, mythical stories are said to be the manifestation of man's sublime and noble ideals and aspirations which do not originate from the science of knowing man but are rooted in the pure and innocent nature of human beings. In fact, myth presents life as it should be not as it is. “Eric Fromm maintains that myth is the symbolic manifestation of philosophical and religious thoughts and beliefs, the occurrence of significant events in time and place, and important psychological experiences of man” (Purnamdarian 1368: 150). Even a philosopher like Plato, who did not have much respect for poets, used myth in his *Republic* in order to project philosophical speculation beyond the point at which certain knowledge is possible (Huby, 1972: 11).

The use of myth allows the poet to give universal significance to his work as it embodies most general human experiences. On the other hand, the symbolic nature of the mythical stories gives potential meanings and interpretations beyond its literal meaning at different times. The potentiality of bearing layers of meaning gives the work everlasting importance. Ferdowsi himself refers to this potentiality when he says:

> Do not suppose that this is mere legends and lies
> They don't have the same meaning over time
> Whatever of it which does not agree with reason
> Signifies symbolic meaning

(*Ferdowsi, 1363; vol. 1: 8*)
Therefore, it is not the case that myth and epic represent unnatural or unreal life, but as Fuller (Bulfin, 1959: 18) argues they represent a type of truth which cannot be revealed unless in the frame of symbolic stories. And, this is the view that we should adopt for reading the Shahnamah; otherwise, we will not be able to enter its realm and understand its wonderful mysterious world. It should be noted here that while mythical stories constitute the first part of the Shahnamah, the use of symbolic language is used throughout the second part as well. Now a question arises here: How can we interpret the mythical and symbolic stories specifically when they do not agree with common sense? Of course, the answer to this question is presented by Ferdowsi himself. He praises man for his reason with the aid of which he is able to understand things. In the introduction to the Shahnamah, he refers to reason as the main basis of creation, thought, faith, existence and life. He talks about reason with such a firm belief that the reader is naturally captivated. The two main virtues which Ferdowsi praises human beings are: his reason and soul. It's for the same reason when he first starts to praise God in the Shahnamah since it is God, who is the creator of reason and soul.

In the name of the Lord of the soul and of wisdom,

than whom thought can conceive nothing higher; the

Lord of all things nameable and of all space; the Lord
Who grants sustenance and is our Guide

(ferdowsi, 1967: 1)

Reason is specifically a human feature which shows itself as soon as he starts to speak. Thus, Ferdowsi chooses reason as the main factor of thought hundreds of years before Descartes and Voltaire. The view Ferdowsi expresses about reason, is, in fact, derived from Islam as the importance of reasoning in dealing with different affairs has been emphasized in many verses of the Quran:

Thus doth Allah make clear His Signs to you:
in order that ye may understand

(quran, Baqare, 242)
Ferdowsi guides his readers to understand his poetry with the help of reason. However, he knows that reason is not enough to discover the essence of everything: for instance, man cannot discover the nature of God with the aid of his reason. Therefore, he gives us another lamp in the light of which we can see things better and this is nothing but symbol. Thus, reason and symbol are the two main components of Ferdowssi's literary theory. This literary theory is also motivated by the Holy Quran while it emphasizes thinking and reasoning, it expresses some truths in the frame of ambiguous verses. Ferdowsi explicitly declares that the stories of *Shahnamah* are true:

> Do not think that this is mere legend and lies.  
> (Ferdowsi, 1363.: vol. 1: 8)

However, there are stories which apparently seem to be legends and pure product of fantasy. It is only the symbolic nature of these stories, which guides us to the depth of their truth. The symbolic stories have developed in the context of Ferdowssi's special Islamic thought and outlook. He uses certain allusions to help us to interpret some of these symbols, for example, in the end of the *Akvane Dive* story, he refers to demons as wicked people who are disobedient towards God:

> Know that evils are wicked people  
> Who are not grateful to God.  
> (Ferdowsi, 1350.: vol. IV: 1058)

Similarly, Rostam is not a mere product of fantasy; rather, he signifies an ideal heroic character that enjoys both physical and spiritual strength. He is never heedless of God. He acknowledges God as the main source of his strength and power; he fights in favor of his God only; and modesty is one of his remarkable virtues. Rostam's greatness is not in his physical capabilities but in the purity and sincerity of his soul and reason. He, in face to face with the Esfandiar, one of the gratest heros of *Shahnameh*, says:
Know that God is my witness
And wisdom is my guide
(Ferdowsi, 1363: vol. IV: 1255)

He symbolizes a spiritual and religious human being in the actual world, the highest model to all human beings, an ideal whose virtues embody the teachings of great religions. He is loyal and devoted to his God and religion, decisive to fight against the carnal soul. He fights against injustice because of his compassion for the weak and resists King Goshtasb when he realizes that Goshtasb is a cruel, oppressive ruler (the same happens when Prometheus stands against the capricious Jupiter (Nodushan, 1356 A.H.: 117).

It is true that the Shahnamah is related to different eras in ancient Iran, but Iran here signifies an ideal country. It is a glorious and magnificent country full of justice and balance, which is the center of goodness, generosity, decency, righteousness, integrity morality, and benevolence. Man, the way Ferdowsi talks about him in the Shahnamah, is also an ideal man who follows his reason not his inhuman passionate desires. He is fair and impartial in his judgments, benevolent and full of love for perfection. These characteristics indicate that Rostam does not belong to a certain geographical region. On the contrary, he stands as a perfect model for all human beings for ever. He is a perfect man who thinks has a mission from God to defend goodness. At the end of the third adventure (khoun), after killing the magic dragon, he praises God and addresses him:

And on the name of God the hero called,
With water washed his body and his head,
To seek the word but in his God was led,
“O just one!” to his maker then he cried
Knowledge thou gavest me and strength and pride.
Be me foes many, or yet be they few,
Grown angry, all but as one man I view.
(ferdowsi, 1973: 145)
Shahnamah is the story of struggle between good and evil. Its main theme is the critical resistance of good against the aggression of the dreadful evil. In this battle Iran is the pivot of goodness and righteousness; and, Turan symbolizes evil. If Iran is praised in Shahnamah, it is not for its own sake but in fact it is the goodness which is praised. Therefore, it is a holy action to defend Iran as the realm of goodness being attacked by the oppressor. The men who appear in the Iranian army endanger their lives for the same supreme cause. Thus, they are also devoted to moral values not racism or nationalism. On the other hand, Turan is the camp of aggression, wickedness, and immorality. The defense of goodness and moral principles is so important that the battle against wickedness becomes a holy one for the heroic figure and which should be respected at any rate. Those who have been innocently murdered cannot be ignored and the sinner who is responsible for their blood must be punished, even though the sinner is an offspring or a close relative of his own. This is why King Fereidun agrees with the death penalty of his son: and King Keykhosro fights against his grand father. Thus, relations, borders, race and nation are all subject to universal moral principles and values.

Similarly, the mythical kings in the Shahnamah stand for different types of people whose history and adventures denote man's struggle against barbarism and savagery and the progress of the civilization (Forughi 1362AH: 30).

In many of the stories of the Shahnamah there is no clear cut geographical border. This point has lead the famous German orientalist, Noldeke to claim that Ferdowsi did not know geography (Bozorge Alavi, 1369AH: 95). But, Noldeke has, in fact, been unaware that the Shahnamah is the story of humanity and mankind as a heroic figure, a story in which all borders except the border between goodness and wickedness lose their importance. The distinction between the good and the bad is so important that issues such as time, place borders, race, tribe, ethnicity, family, all seem too trivial to deal with. This characteristic has made the Shahnamah an everlasting and universal masterpiece which stays fresh and vivid in all times and ages being read effectively. Noldeke admits that the Shahnamah is an eternally unique masterpiece which no literary work will ever be equal. A masterpiece which teaches us how to live and how to die.
Thus, in creating the *Shahnamah*, Ferdowsi has concentrated on principles and norms which make the very essence of humanity. The humanistic values motivate action all throughout the *Shahnamah* as it is the case when the biggest battles occur in order to defend Siavash and his innocence. It should be mentioned here that these principles and norms are the same values which form the dominant themes in Islamic religious texts. As mentioned above, most of these themes and concepts are implicated symbolic in the *Shahnamah*, and an inability to realize their mythical and symbolic aspects would lead to the ignorance of truths they are aimed to convey.

**Ferdowsi and his faith**

There is a misunderstanding about Ferdowsi and that is the belief that the ancient religions of Iran are the origin of his beliefs. Berteles, the Russian orientalist (1362AH: 190), without presenting any evidence and based on a superficial analysis, claims that dualism is the essence of Ferdowsi's beliefs. Zaehner (1955), an Oxford scholar asserts that Ferdowsi's religion was Zorvanism whose ideas were popular during the Sasani era. However, the famous researcher of the *Shahnamah*, Noldeke (1369AH: 72) rejects the idea that Ferdowsi has been a dualist or a follower of Mazdism and states that Ferdowsi had certainly been a monotheist. In fact, a close scrutiny of the *Shahnamah* and specifically its introduction clarifies the point that it is only Islamic faith which has inspired him to design the greatest epic of the world. Surely, there are many facts about the *Shahnamah* which indicate that nothing frames the core of Ferdowsi's beliefs except Islamic ideas.

First, he has been aware of the Islamic traditions and history, the translation of which can be found here and there in the *Shahnamah* (Foruzanfar, 1369AH: 48).

> First of creation do thou wisdom see,<br> The soul’s own guardian, prayers’s own guard to thee,<br> Thy prayer to thee is eye and ear and speech,<br> Through these both good and evil thee reach.<br> (Ferdowsi, 1973: 3)
The translation of this Hadith:

The first Thing that God created is wisdom

(Foruzanfar, 1361: 202)

Second, his Islamic faith is evidently expressed in his admiration of the Prophet and Islamic leaders in the introduction to the Shahnamah. Moreover, many verses in the Shahnamah have been devoted to the Prophet (peace be on him) indicating the depth of Ferdowsi belief and his perfect faith in what he declares.

Indeed the knowledge and religion rescue you
The path of salvation must be investigated
Find the sayings of the prophet (Muhammad)
And wash away the darkness from your heart
I am the follower of the prophet
The worshiper of the soil under the feet of his successor(Ali-ben- Abitoaleb)

(Ferdowsi, 1363: vol. 1: 62)

Third, although Ferdowsi has admired Zoroaster and his religion, he has detestably rejected dualism and the worship of fire which are the core of Zoroastrian faith. He even claims that the essence of Zoroastrainsim is clear from polytheism or dualism, and constantly praises the unique God in the words of kings and heroes. In a letter to Zal that represented of Rostam, he described his victory over the foe in this way:

First praise to God who created sun
Who created snakes an ants
Who created Mars, Venuse and sun
Who created the high sky

(Ferdowsi, 1363: vol. 1: 182)

Fourth, his view that God created the Thing from Nothingness is an Islamic idea which sharply contrasts with Zoroastrianism (Rajaee 1356AH: 458).
For God from nothing every thing has made,
So that His power might be thus displayed.
(ferdowsi, 1973: 3)

Fifth Ferdowsi respects other religions with great consideration, but his own
religion is undoubtedly Islam. His belief in the unity of God, prophethood,
resurrection, eternity of the soul, paradise and hell, which are all main elements of
Islamic faith, are evidently manifested in different parts of the *Shahnahah*.

I have believed this idea and believe it for ever
Surely, I am the follower of Heidar(Ali-ben- Abitoaleb)
(Ferdowsi, 1363: vol. 1: 8)

Egocentrism or self-indulgence, which is widely blamed and rejected in Islamic
texts, is the source of all the wickedness and rebels in the *Shahnahah*. It is self-
conceit which brings about the decline of Jamshid's throne usurped by the brutal
Zahhak whose reign brought a thousand years of slavery and suffering for Iranians.
Jamshid to the sages of his tribe arrogantly said:

None in the word I know myself besid.
Only through me the word has virtue known:
Like me has none possessed the royale throne.
The word in beauty I alone arrayed,
All troubl on the face of earth allayed
(ferdowsi, 1973: 20)

And this arrogance spoiled Iranian peace and brought about disputes and wars
and culminated to the end of Zahhak a thousand year ruling:

There rose on this from all Iran a shout,
On all sides war and tumult raged about,
The white and brilliant day to darkness turned,
And men of Jamshid the connection spurned…
When Zahhak sat upon the throne asking,
A thousand years of him were on the wing.
Fortuon to him displayed on evil face,
But yet he lived a very lengthy space.
The custums of the good were then concealed,
And of the mad ones evry hope revealed.
Virtue was scorned and majic had its way,
Hidden was truth, ill practised in the day.
The Divs streched out their hands towrds ill alone,
Man spoke of good But in a lowered tone.II

(ferdowsi, 1973: 27&30)

Another recurrent theme in Shahnamah is the unreliability of the world which is unquestionably one of the topics in Islamic thought. This unreliability is evidenced by the death phenomenon. Death and mortality are basic elements to remind man of the outcome of his deeds. No one can rely on his endowments and worldly powers which cannot support him for ever. Even Isfandiar who is blessed with the sacred water and has become immortal cannot avoid death. Rostam, the most powerful of all, who does not surrender to any power, is also helpless before death and subject to his fate, meets death, the end point of every life. In the begining of the story of the Rostam and sohrab, Ferdowsi talks about the cruelly of death and its mystry:

If the death is justice, what is injustice
Why is much cry and shout about justice
You may not understand this secret
You can not find this way

(Ferdowsi, 1363: vol. 1: 345)

Sixth, most of the instructions and teachings in the Quran are on the basis of thee principles of unity, resurrection and good conduct. All the prophets have been sent to establish these three words. A close examination of the Shahnamah indicates
that the theme of its stories is nothing but the illustration and clarification of these three ideas. All the stories are related to the unity of God one way or another. All the heroes entering the battlefield start their fight with God's praise. All through *Shahnamah* death and resurrection is mentioned; and finally, all the behavior and deeds of people are judged by the standards of good conduct.

Seventh, when Ferdowsi wants to speak about the value of verbal art, he states that the best example of speech is Quran, a miracle which no man-made speech can ever surpass:

> If God's word was not good
> How could Quran guide us?
> (Ferdowsi, 1363: vol. 1: 23)

Eighth, whatever Ferdowsi comments on, for example concepts such as reason, greed, hypocrisy, oppression, justice, and the like are evidently compatible with the Quran and Islamic tradition, rather than any other school of thought (Farzam 1990). Ferdowsi acknowledges that Islam is the best and greatest religion and explicitly talks about his faith in Islam and its noble leaders. Of course, Ferdowsi is not a mystic lover of Islam but he must be considered as a benevolent well-wisher of goodness and good conduct who deeply beliefs in Islam and its principles (Reza 1356, A.H. : 458 and Nodushan, 1363, A.H. :2). He devotes the opening and closing verses of the *Shahnamah* to the praise of God and his unity, which is also repeatedly mentioned all throughout the epic. He believes that paradise and hell are real truths; and when people leave this world they will be present before God. His description of the paradise is the same as that of the Quaran:

> (Here is) a parable of the Garden which the righteous are promised:
in it are rivers of water incorruptible; rivers of milk of which the taste never changes; rivers of wine, a joy to those who drink; and rivers of honey: pure and clear…
> (Quarun, mohammad, 15)
The God, the owner of the rivers of wine and honey
The same, the spring of milk and clear water
(Ferdowsi, 1363: vol. 1: 6)

The hereafter is where man's deeds will be judged. He remembers Kaba which has been built by Abraham by the command of God with great respect.

Other ideas which are observed in the *Shahnamah* and clearly indicate Ferdowsi's deep knowledge and belief in Islam are his comments on other prophets such as Xezr, Abraham and Moses, who are all mentioned in Quran. His belief that Jesus was the messenger of God and that he was not crucified; the idea that whoever is killed in an Islamic war is a martyr (Shirani1374A.H:200); the superiority of reason and speech to other human merits, the dignity of the artists; love toward God and people and the importance of serving and helping them are all lessons which he has learned from Islam.

Ninth, although Ferdowsi has referred to Greek and Iranian ideas about the creation of the universe, he is more affected by the Islamic views on the basis of which he has judged what other schools of thought have claimed about issues such as genesis of the world, the creation of the earth, heavens and stars (Mohajerani 1372A.H.).

Therefore, on the basis of these observations, it is not unreasonable to call Ferdowsi 'Hakim' because in Islamic culture Hakim refers to a man of philosophy and theology. And, In 'Sharia', the Islamic code, Hakim refers to someone whose words and action are compatible with the rules of religion (Jorjani, 1345A.H: 182).

Another Islamic aspect observed in *Shahnamah* is the dissemination of moral values. Ferdowsi, who has been deeply impressed by the Islamic culture and has been aware of the moral teachings of Quran and Islamic traditions has most eloquently manifested them in poetic language. Numerous moral verses in *Shahnamah* indicate that it has been extremely intended to be a didactic literary work and that Ferdowsi has been delicately instructing people that it is only good deeds which sanctions man's salvation.
Horace in his *Ars Poetica* provides the classical formula for defining literature when he states: “...that poets teach and delight” (Danziger & Johnson 1968). Ferdowsi both delights and instructs at his best combining the useful and the beautiful. In order to achieve this goal he brings up concepts such as goodness and wickedness, truth and lie, honesty and infidelity, right and wrong, justice and injustice, bashfulness and impudence, kindness and cruelty, forgiveness and revenge, among many others; and thus, creates the loftiest ethical teachings based on the Islamic worldview in his poetry. That is why chastity and purity dominates the atmosphere of the stories in *Shahnamah*. The women in *Shahnamah* with the exception of Sudabeh are all among the most pious and virtuous; and their love is always accompanied with chastity.

**Conclusion**

Considering different aspects of *Shahnamah* including its themes, the concepts involved, and its symbolism indicate that the scope of this literary work is not confined to the borders of a certain country or race. Rather, what we find is the fact that it aims at a higher and more supreme goal. It addresses the human race seeking peace, unity, and perfection for everyone. And thus, it achieves a universal function which mirrors Islamic thought and purpose.

More precisely, with his peculiar outlook, Ferdowsi vividly pictures the human race with all his positive and negative potentials, his strengths and weaknesses and his opposing motives. Then taking a wide scope considers man’s position in the universe and emphasizes his eternal peculiarities in order to show him his ultimate perfection. Thus, he creates a sensational epic with the most elegant and harmonious manifestation of heroic, literary, historical, intellectual, philosophical, and religious concepts and issues. It is true that *Shahnamah* is displaying Persian pride and character as a nation and the essence of their civilization and culture. But, at the same time, it relates the heroic action of a dynamic mankind, who by relying on spiritual values fights against his mighty fate in order to prove himself as the peak of excellency and perfection. In Ernest Renan's words, Ferdowsi believes in honor and
spirituality. He is a humanist who acknowledges civilization as the optimal purpose of man's existence (Sajjadi 135A.H.: 10). These are the reasons which explain why the *Shahnamah* has achieved widespread popularity all over the world as an Islamic art.

References


Quran. Translated Elahi Qoumshei, Islamic Organization Qoum.


Formulaic Writing: A Novel Approach to Writing Instruction

Behzad Ghonsooly Hazare*
Professor of English department, Ferdowsi university of Mashhad, I.R. Iran

M. Saber Khaghaninezhad**
Ph.D student of TEFL, Ferdowsi university of Mashhad, I.R. Iran

Hesam-al-din Shahriari Ahmadi***
Ph.D student of TEFL, Ferdowsi university of Mashhad, I.R. Iran

(Received: 12 July. 2009, Accepted: 22 Nov. 2009)

Abstract
The present study seeks to introduce and investigate an approach to teaching writing in the EFL classroom which draws heavily on the use of formulaic expressions. In contrast to its predecessors, this approach includes both the application of process writing techniques and the teaching of grammar and vocabulary into its syllabus. Two groups of 30 EFL learners participated in this study. One group was taught writing through the proposed method, while the other group received instruction of a grammatical and structural nature. Following the courses, learners of both groups were asked to write an argumentative essay. A comparison of the essays written by the participants of the two groups revealed that those who were taught with the formulaic approach were more successful in fulfilling the purpose of the task, as well as creating more coherent and structurally accurate texts. The authors of this study argue for the use of formulaic expressions in teaching writing, especially in proficiency test preparation courses, due to its usefulness in fulfilling the communicative requirements of such examinations.

Key Words: Formulaic Expressions, Formulaic Writing, Process Writing, Argumentative Essay.

* Tel: 0711-6244047, Fax: 0711-6244047, E-mail:ghonsooly@yahoo.com
** Tel: 0711-6244047, Fax: 0711-6244047, E-mail: saber.khagham@yahoo.com
*** Tel: 0711-6244047, Fax: 0711-6244047, E-mail: hesamshahriari@yahoo.com
1. Introduction

Researchers and teachers increasingly acknowledge the need of eliciting samples of language use which are representative of learners’ performances in real-world communications when accuracy is not the main focus. Such samples, it is believed, provide evidence of learners’ abilities to use their second language knowledge in real-life contexts of meaning negotiation. Believing in the importance of such samples for documenting how learners structure and restructure their interlanguages over time, second language acquisition (SLA) researchers recognized that unless learners are given opportunities to experience such samples they may not succeed in developing the required language proficiency they need to communicate fluently.

Having this at the back of their minds, researchers/teachers tried to embody their aims through using task-based language teaching and learning (TBLT) techniques mainly for two reasons; one was their commitments to a form of teaching that treats language primarily as a tool of communication rather than an object for study or manipulation. Clearly, if learners are to develop the competence they need to use a foreign language easily and effectively in unexpected situations they encounter outside the classroom, they need to experience how language is used as a tool for negotiating meaning and the second, was their wishes to see how second language acquisition develops not just as an autonomous discipline (it seems that it clearly has moved in this direction during the last decade) but also as an applied area of study. According to Ellis (2003), usage of formulaic expressions in writing instruction is one of the significant factors which revolutionize current methods of teaching writing and change them into more real-life trends.

The present study seeks to investigate the effect of an approach to teaching writing in EFL classes which draws heavily on the use of formulaic expressions by comparing the final results of learners who were taught under formulaic approach (they were provided with eight formulaic structures and were made aware of the purposes they serve) with those who were taught writing by traditional approach of writing instruction (instruction of language features and structural elements) and
tries to answer the following questions:

- Does usage of formulaic expressions in writing instruction have any positive effects on writing proficiency improvement of language learners?
- Which aspects of learners’ writing proficiency would be improved under formulaic instruction?

2- Literature review

2-1 Writing tasks classification

It seems imperative to have a brief look at writing tasks typology prior to referring to the type of writing instruction developed through the paper. Here, some world-widely accepted categorizations toward various writing task types have been presented. Broadly speaking, each writing task can be put into one of these proposed categories;

Rod Ellis (2003)

According to Ellis, the survey of the research literature on writing tasks reveals a bewildering variously-labeled array of task types. There are different ‘gap’ tasks, for example, information-gap and opinion-tasks, which are also sometimes referred to in terms of how the information has been organized in the task, i.e. split versus shared information tasks. There are also reciprocal and non-reciprocal writing tasks, i.e. tasks that require or do not require interaction to achieve the outcome. Writing tasks can be labeled according to the kind of activity they require of the learner or according to the language skills they focus on. They can be named according to the type of discourse they are intended to elicit, for example narrative or descriptive tasks.

Bley-Vroman (1993)

He believes in *structure-based writing tasks* and in key article discusses what he calls’ structure-based communication task’. He distinguishes three ways in which
a task can be designed to incorporate specific target language feature. The first is task-naturalness. In this case, the target structure may not be necessary for completion of task but nevertheless can be expected to arise naturally and frequently in performing the task. The example Bley-Vroman gives is of a task that involves the exchange of information about a travel itinerary. He suggests that this will lead naturally to the use of the present tense. He refers to research on interlanguage variation, which has shown that different types of tasks result in different uses of grammatical features.

The second way of incorporating a linguistic focus is in terms of task-utility. By this Bley-Vroman means that even though a targeted feature is not essential for completing the task it is very useful in performing other tasks. Of course, as he acknowledges, the utility of a structure is relative to the learner’s existing stage of acquisition. He points out that students who have already achieved full mastery of a specific structure will not benefit acquisitionally from producing the structure.

The third way of designing a focused task is to try to ensure the task-essentialness of the targeted task. This requires that learners must use the feature in order to complete the task successfully. In this respect, the targeted feature becomes the essence of the task. However, the examples Bley-Vroman gives are all comprehension rather production tasks. He acknowledges that it may be impossible to design writing tasks that make the production of the target feature essential and, in fact, task-essentialness can only be achieved by receptive tasks.

Barbara Kroll (1996)

Kroll mentions that advanced writing activities shift their goal from the focus on the mechanics of writing to basis process-oriented tasks which need to incorporate some language work at the morphological and discourse level. Thus, these activities will enable focus on both accuracy and content of the message. She continues, in order to develop and use these more demanding writing activities in the ESL/EFL classroom, we need to develop a detailed set of specifications which will enable both teachers and students to cope successfully with these tasks. Such a set of
specifications should include the following:

- **Task description**: to present students with the goal of the task and its importance. **Content description**: to present students with possible content areas that might be relevant to the task.

- **Audience description**: to guide students in developing an understanding of the intended audience, their background, needs, and expectations.

- **Format cues**: to help students in planning the overall organizational structure of the written product.

- **Linguistic cues**: to help students make use of certain grammatical structures and vocabulary choices.

- **Spelling and punctuation cues**: to help students focus their attention on spelling rules which they have learned and eventually on the need to use the dictionary for checking accuracy of spelling, and to guide students to use acceptable punctuations and capitalization conventions.

**Elite Olshtain (1991)**

He divides writing tasks into three categories:

**Practical writing tasks**: These are writing tasks which are procedural in nature and have a predictable format. This makes them particularly suitable for writing activities that focus primarily on spelling and morphology. Lists of various types, notes, short messages, simple instructions, and other such writing tasks are particularly useful in reinforcing classroom work. Lists can be of many types; “things to do” lists, and “things to complete” lists, or shopping lists. Each of these list types provides us with an opportunity to combine some spelling rules with morphological rules and with the logical creation of a meaningful message.

**Emotive writing tasks**: Emotive writing tasks are concerned with personal writing. Such personal writing primarily includes letters to friends and narratives describing personal experiences, as well as personal journals and diaries. When dealing with letter writing, emphasis can be placed on format, punctuation, and spelling of appropriate phrases and expressions. When writing about personal
experiences usually done in a narrative format, spelling of past-tense forms can be reviewed and practiced. It seems that emotive writing, to serve the personal needs of the learners, has to be quite fluent. The different types of emotive writing activities are, of course, suitable for the more advanced courses, but they can be carried out, in a more limited manner, even at the initial stages. Thus, personal letters can be limited to the level of structural and vocabulary knowledge of the students at each point of time. Similarly, journal and personal writing activities can reflect the learner’s proficiency level.

School-oriented tasks: One of the most important functions of writing in a student’s life is the function it plays at school or university. It is still the case that much individual learning goes on while students are writing assignments, summaries, answers to questions, or a variety of essay-type passages.

In most case, the audience for these writing tasks is the teacher, but gradually students must learn to write to an unknown reader who needs to get information being imparted exclusively via writing. It is the combination of the content and organization with accepted formal features that will lead learners to better utilization of the writing skill in their future use of English.

2-2 Approaches to instruct writing

Nold (1982) proposed four general approaches for preparing students to meet their writing needs. To him, intermediate- and advanced-level ESL/EFL writing courses generally have one of four orientations, depending on which element of composing is taken as the basis for course organization;

Rhetorical approaches ask students to analyze and practice a variety of rhetorical or organizational patterns commonly found in academic discourse: process analysis, partition and classification, comparison/contrast, cause-and-effect analysis, pro-and-con argument, and so on. Kaplan (1966) and others point out that rhetorical patterns vary among cultures and suggest that nonnative students need to learn certain principles for developing and organizing ideas in American academic discourse, such as supporting generalizations by presenting evidence in inductive
and deductive patterns of arrangement.

*Functional approaches* recognize that in real writing, purpose, content, and audience determine rhetorical patterns. Starting from given patterns and asking students to find topics and produce essays to fit them is thus a reversal of the normal writing process. Instead of having students write a comparison/contrast essay, a functional approach would ask students to start with a specified purpose and audience, for example, “Persuade one of your friends who is planning to move that City X is a better place to live than City Y.” Typically, in a functionally oriented writing program, writers assume a variety of roles; academic writing is only one context and usually not the sole focus. Contexts for writing tasks are carefully defined; purpose and audience are always specified. If the writer is placed in unfamiliar roles in which background knowledge about the topic may be lacking, data may be supplied in the form of facts, notes, tables or figures, quotations, documents, and so on.

*Process-centered approaches* help student writers to understand their own composing process and to build their repertoires of strategies for prewriting (gathering, exploring, and organizing raw material), drafting (structuring ideas into a piece of linear discourse), and rewriting (revising, editing, and proofreading). Tasks may be defined around rhetorical patterns or rhetorical problems (purpose), but the central focus of instruction is the process leading to the final written product. Students are given sufficient time to write and rewrite, to discover what they want to say, and to consider intervening feedback from instructor and peers as they attempt to bring expression closer and closer to intention in successive drafts.

*Content-based approaches* differ from traditional approaches to teaching academic writing in at least four major ways:

1. Writing from personal experience and observation of immediate surroundings is de-emphasized; instead, the emphasis is on writing from sources (readings, lectures, discussions, etc.), on synthesis and interpretation of information currently being studied in depth. Writing is linked to ongoing study of specific subject matter in one or more academic disciplines and is viewed as a means to stimulate students
to think and learn.

2. The focus is on what is said more than on how it is said (Krashen, 1982) in preparing students for writing and in responding to writing. The instructor who guides and responds to writing must know the subject matter well enough to explain it, field questions, and respond to content and reasoning in papers. Treatment of matters of form (organization, grammar, and mechanics) and style do not dictate the composition course syllabus, but rather follow from writers’ needs.

3. Skills are integrated as in university course work: Students listen, discuss, and read about a topic before writing about it—as contrasted to the traditional belief that in a writing course, students should only write.

4. Extended study of a topic (some class treatment of core material and some independent and/or collaborative study/research) precedes writing, so that there is “active control of ideas” and “extensive processing of new information” (Anthony, 1985) before students begin to write. A longer incubation period is permitted, with more input from external sources, than in traditional composition classes, in which students rely solely or primarily on self-generated ideas and write on a new topic for each composition.

2-3 Formulaic expressions

There are certain pre-coded (formulaic) utterances conventionally triggered by certain communication situations, and their use is expected and deemed appropriate because they are seen as part of everyday politeness formulas (they are also called phatic structures in the literature). Some routines are taught explicitly and their use is prompted by adults in socializing children. These expressions are part of every competent speaker's repertoire, and include proverbs, idioms, greetings, apologies, thanks, and leave-taking. Human language is distinguished by its creative potential. New sentences, never spoken or heard before, can easily be formulated given the set of rules for combining a large set of vocabulary items (Chomsky, 1965). As linguists have explained for many years, this system allows for the generation of an infinite set of context-free sentences from a finite grammar. The standard view in linguistic
textbooks is stated by Pinker (1995);

...virtually every sentence that a person utters or understands is a brand-new combination of words, appearing for the first time in the history of the universe (p.22).

This statement, however, is seriously misleading. Many utterances in everyday language are conventional expressions that must be used in a certain way. Conventional or formulaic expressions (FEs) are distinguished from novel utterances in a number of ways. They often contain lexical items with non-literal or nonstandard meanings. Unlike novel sentences, which can be strictly neutral in affective content, FEs are generally laced with attitudinal or emotional innuendoes. FEs are “familiar” in the sense that a native speaker will recognize them as having a special status. As stated by Jackendoff (1995), a very large number of a broad range of formulaic expressions “are familiar to American speakers of English; that is, an American speaker must have them stored in memory”. It follows that a survey using recall and recognition tasks adapted for the study of FEs can provide objective and quantifiable data to support the claim that native speakers “know” FEs. In English speaking cultures, not only are FEs often subsumed under the opprobrious label of “cliches”, but also current linguistic models emphasize combinatorial creativity as the central property of human language (Van Lancker, 2001).

2-4 The importance of language chunks in writing proficiency

Several publications in the past 15 years have highlighted the importance of formulaic language chunks (i.e., multiword phrases and routines treated as single lexical units) in both L1 and L2 use. Although these chunks are variously referred to by different authors as gambits (Emig 1977), conventionalized language forms (Murray 1985), lexical phrases (Newell 1984), conversational routines, pre-patterned speech (wilkinson 1985), lexicalized sentence stems (Zamel 1983), partially pre-assembled patterns (Widdowson, 1990), or formulaic constructions (Pawley, 1992), all these authors agree that such chunks play a more significant role
in language production than is normally acknowledged.

Native speakers of a language are in command of thousands of language chunks and use them as building blocks in their writings. The retrieval of these chunks is cognitively relatively undemanding, which allows the writer to attend to other aspects of communication and to plan larger pieces of discourse. L2 learners, on the other hand, often put sentences together from scratch, that is, word by word, which takes up their cognitive capacity and does not let them achieve native-like fluency. Nattinger and DeCarrico (1992) discuss in detail how lexical phrases can serve as an effective basis for a new, increasingly lexis-oriented teaching of writing, and, indeed, there have been indications in L2 methodology that such a development is more than a mere theoretical possibility.

Given the assumed importance of formulaic language chunks, it is surprising that, until the recent publication of Nattinger and DeCarrico’s *Lexical Phrases and Language Teaching*, no comprehensive study had been written offering a systematic and empirically based analysis of the issue. Nattinger and DeCarrico’s high-quality work in *Lexical Phrases and Language Teaching*, coupled with the fact that it has filled such a noticeable and long-existing gap in applied linguistics, led to the books being awarded the Duke of Edinburgh Prize in 1992, the foremost British award in TESOL and applied linguistics. *Lexical Phrases and Language Teaching* is a classic applied linguistic work as it contains a thorough linguistic introduction to and analysis of the issue as well as a detailed discussion of how the teaching of foreign/second languages can benefit from the theoretical insights. The authors base their arguments on a review of the existing literature, analysis of a broad corpus of spoken and written English, and their own experience in the teaching of lexis.

3- Method

The participants of this study consisted of 60 (36F, 24M) learners of English as a foreign language studying at *Gaame Andisheh Shargh* language institute in the city of Mashhad, Iran. The students included in this study were all taking part in General English and IELTS preparation courses and trying to improve their writing skills to
satisfy their academic future needs at the time of selection. The mean age for the participants was 23, with the youngest and oldest participants being 19 and 42, respectively.

Participants were divided into two groups of 30, based on their results on a language proficiency test. The language proficiency test used for this purpose was the paper-based version of the Test of English as a Foreign Language (TOEFL). This test was comprised of listening, grammar and reading sub-sections. Despite being regarded as a stand-alone sub-test, the Test of Written English (TWE) was also included in the pre-test. The paper-based TOEFL consists of 140 multiple choice items and the TWE is scored on a scale of 6 points. The results obtained by the two groups on the listening, grammar and reading sections of the pretest can be seen in figure 1.

<table>
<thead>
<tr>
<th>GROUPS</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>30</td>
<td>116.76</td>
<td>10.77</td>
<td>1.96</td>
</tr>
<tr>
<td>2.00</td>
<td>30</td>
<td>115.36</td>
<td>11.90</td>
<td>2.17</td>
</tr>
</tbody>
</table>

Figure 1. Descriptive statistics for the TOEFL

In order to determine whether there was a significant difference between the means obtained by the two groups on the TOEFL, the independent samples t-test was used. The results of this test revealed that the two groups did not significantly differ from each other in terms of their performance on the TOEFL. Figure 2 shows the results of the independent t-test.
Levene's Test for Equality of Variances

<table>
<thead>
<tr>
<th></th>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOEFL</td>
<td>.54</td>
<td>.46</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>.47</td>
<td>57.43</td>
</tr>
<tr>
<td>TOEFL</td>
<td>.47</td>
<td>.46</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>.47</td>
<td>57.43</td>
</tr>
</tbody>
</table>

Figure 2. Independent samples t-test for the TOEFL

The TWE was also of great relevance to the division of learners into two groups. That is, it was of particular importance to ensure the equality of the two groups in terms of their writing proficiency in English. Each writing performance was independently marked by two raters, each of whom underwent a program of rater training (i.e., a norming session), consisting of an orientation to the writing test and a tutorial involving a number of sample responses.

Both raters were graduate students of TEFL and were experienced teachers of English as a foreign language. It is worth mentioning that the same raters also cooperated in the scoring of the post-test. The independent t-test was also used to determine whether the two groups exhibited any significant difference with regards to their writing proficiency in English. The descriptive statistics and the results of the independent t-test can be seen in figures 3 and 4, respectively.

Figure 3. Descriptive statistics for TWE

<table>
<thead>
<tr>
<th>GROUPS</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWE</td>
<td>1.00</td>
<td>30</td>
<td>3.50</td>
<td>.93</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>30</td>
<td>3.43</td>
<td>.89</td>
</tr>
</tbody>
</table>
As can be seen in figure 4, there was no significant difference in the mean scores obtained by the participants of the two groups on the TWE. Considering the results of the pre-tests, it could be claimed that the participants of the two groups were equal with regards to their overall language ability, as well as their writing proficiency in the English language.

The learners of each group participated in an EFL writing course intended to learn how to write an argumentative essay. Both groups underwent 18 hours of instruction. The classes were held twice a week at Gaame Andisheh Shargh language institute, with each session lasting for about 90 minutes. Both classes were taught by one of the researchers who was an experienced EFL instructor and a regular teacher at the institute. The learners were informed that the course was carried out as part of a research project and were charged a nominal fee to take part in the course. However, learners were kept unaware of the aims of the research.

The first group participated in a course, in which the focus was chiefly on the instruction of language features and structural elements of the English language. The learners were familiarized with various types of phrases and clauses of the English

<table>
<thead>
<tr>
<th>Equal variances assumed</th>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>.12</td>
<td>.72</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>.28</td>
<td>57.89</td>
</tr>
</tbody>
</table>

**Figure 4. Independent samples t-test for TWE**
language, and were provided with relevant activities and exercises to the structures being taught each session. In the final session of the course, learners were given the chance to use the materials covered in the previous sessions to practice writing an argumentative essay. All essays were corrected by the teacher, who also provided comments on the learners’ use of language. The syllabus for the first course has been summarized in table 1.

<table>
<thead>
<tr>
<th>Session</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction to sentences</td>
</tr>
<tr>
<td>2</td>
<td>Introduction to clauses, phrases and conjunctions</td>
</tr>
<tr>
<td>3</td>
<td>Adverbial clauses</td>
</tr>
<tr>
<td>4</td>
<td>Adjective clauses</td>
</tr>
<tr>
<td>5</td>
<td>Noun clauses</td>
</tr>
<tr>
<td>6</td>
<td>Participle phrases</td>
</tr>
<tr>
<td>7</td>
<td>Gerund phrases</td>
</tr>
<tr>
<td>8</td>
<td>Infinitive phrases</td>
</tr>
<tr>
<td>9</td>
<td>Absolute constructions</td>
</tr>
<tr>
<td>10</td>
<td>Abstract noun phrases</td>
</tr>
<tr>
<td>11</td>
<td>Appositive phrases</td>
</tr>
<tr>
<td>12</td>
<td>Writing Practice and feedback</td>
</tr>
</tbody>
</table>

Table 1. Syllabus for the course delivered to the first group of participants

The second group of participants underwent treatment in the form of a 12-session course which focused on teaching writing proficiency through the presentation of fixed formulaic expressions. This course differed from the first course in that it did not attend to language features of the English language, except for brief explanations on the difference between phrases and clauses, various conjunctions and also the use of present participle phrases. Contrary to the first course, learners of the second group were taught how to organize and prepare their
ideas for use in the formulaic structures which were later presented to them.

In the second course, learners were provided with eight formulaic structures and were made aware of the purposes they serve. Two structures were used in the introductory paragraph. Three structures were introduced to serve as the topic sentence of each body paragraph. Two structures served the purpose of providing context for supporting ideas and finally, one structure was taught to be used in the concluding paragraph. The syllabus for the treatment course has been presented in table 2 below.

<table>
<thead>
<tr>
<th>Session</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction to the aims and features of argumentative writing</td>
</tr>
<tr>
<td>2</td>
<td>Brainstorming</td>
</tr>
<tr>
<td>3</td>
<td>Introduction to main ideas and supporting ideas</td>
</tr>
<tr>
<td>4</td>
<td>Organizing and planning paragraphs</td>
</tr>
<tr>
<td>5</td>
<td>First structure for the introductory paragraph</td>
</tr>
<tr>
<td>6</td>
<td>Second structure for the introductory paragraph</td>
</tr>
<tr>
<td>7</td>
<td>Structures for topic sentences 1</td>
</tr>
<tr>
<td>8</td>
<td>Structures for topic sentences 2</td>
</tr>
<tr>
<td>9</td>
<td>Structures for supporting ideas 1</td>
</tr>
<tr>
<td>10</td>
<td>Structures for supporting ideas 2</td>
</tr>
<tr>
<td>11</td>
<td>Structures for the concluding paragraph</td>
</tr>
<tr>
<td>12</td>
<td>Writing practice and feedback</td>
</tr>
</tbody>
</table>

Table2. Syllabus for the course delivered to the second group of participants

Each structure was comprised of a series of fixed phrases, from which the learners could choose to create variety in their writing. The structures also had some blank spaces, which had to be filled by either a phrase or a simple clause by the learners. The formulaic structure for the first sentence of the introduction paragraph has been shown in figure 5.
At the end of each course, learners were given the second task of the International English Language Testing System (IELTS) to complete. For this task, participants were given an argument, for which they had to write a 250 word essay, supporting their position and refuting the opposing side. The time allocated to the completion of this task was 40 minutes. The same task was presented for both groups. The participants of both groups took part in the post-test on the same time and date.

4- Data analysis

The post-tests were scored by the same raters who rated the pretests. The essays were marked on a scale of 1-9 for their coherence, cohesion, fulfillment of purpose, structural variation and lexical variation (the most needed writing elements for their future academic ventures). The results of the independent sample t-test for each of the criteria were computed. Figures 6 to 10 reveal the results of the t-test for different criteria in the writing of the two groups.
### Levene's Test for Equality of Variances

<table>
<thead>
<tr>
<th>Coherence</th>
<th>Equal variances assumed</th>
<th>Equal variances not assumed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>.55</td>
</tr>
</tbody>
</table>

**Figure 6. Independent samples t-test for coherence on the post-test essay**

### t-test for Equality of Means

<table>
<thead>
<tr>
<th>Coherence</th>
<th>Equal variances assumed</th>
<th>Equal variances not assumed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>.04</td>
</tr>
</tbody>
</table>

**Figure 7. Independent samples t-test for cohesion on the post-test essay**
<table>
<thead>
<tr>
<th></th>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
<td>df</td>
</tr>
<tr>
<td>Fulfillment of purpose</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>.38</td>
<td>.53</td>
<td>4.46</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td></td>
<td></td>
<td>4.46</td>
</tr>
</tbody>
</table>

Figure 8. Independent samples t-test for fulfilling the purpose of task on the post-test essay

<table>
<thead>
<tr>
<th>Structural Variation</th>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural Variation</td>
<td>F</td>
<td>Sig.</td>
<td>df</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>.42</td>
<td>.51</td>
<td>4.53</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td></td>
<td></td>
<td>4.53</td>
</tr>
</tbody>
</table>

Figure 9. Independent samples t-test for structural variation on the post-test essay
Table showing the results of Levene's test for equality of variances and t-test for equality of means.

<table>
<thead>
<tr>
<th>Equal variances assumed</th>
<th>Equal variances not assumed</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>.45</td>
</tr>
<tr>
<td>Sig.</td>
<td>.50</td>
</tr>
<tr>
<td>t</td>
<td>-.72</td>
</tr>
<tr>
<td>df</td>
<td>58</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.47</td>
</tr>
<tr>
<td>Mean Difference</td>
<td>-.23</td>
</tr>
<tr>
<td>Std. Error Difference</td>
<td>.32</td>
</tr>
<tr>
<td>95% Confidence Interval of the Difference</td>
<td></td>
</tr>
<tr>
<td>Lower</td>
<td>-.87</td>
</tr>
<tr>
<td>Upper</td>
<td>.41</td>
</tr>
</tbody>
</table>

Figure 9. Independent samples t-test for lexical variation on the post-test essay

The figures reveal that the essays written by the participants of the two groups were significantly different in terms of coherence, fulfillment of purpose and structural accuracy. Nevertheless, no meaningful difference was observed in the cohesion and lexical variation of the essays written by the two groups. In all cases in which significant differences were reported, essays written by the participants who had undergone treatment (i.e., participants of the experimental group) were rated as being better than those of their counterparts. In other words, the essays by learners in the second group were more coherent, better fulfilled the purpose of the task and exhibited more structural variation. The sample introductory paragraphs below illustrate the differences discussed. The first and second paragraphs were written by advanced participants of the first and second groups, respectively.

People use their personal cars for transportation. They use it even for short distances. This causes many problems including air pollution, high traffic and etc. Governments should introduce laws to restrict the use of personal cars. This can be done in many ways. Some of these ways have been discussed in this essay.

The question of whether or not international laws should be introduced to control car ownership has aroused many debates in recent decades. While some
believe that car use should be controlled, due to social, environmental and political reasons; many hold the belief that the use of public transportation is not always possible, especially in developing countries.

The two paragraphs are approximately of the same length. Neither the first nor the second paragraph suffers from severe grammatical errors. However, the most evident difference between the two lies in their coherence, and more importantly, the extent to which they serve their purpose as introduction paragraphs.

The first writer starts by exploring the current situation and then goes on to explain his point of view and provide some suggestions. The paragraph written by the first writer does not appear to belong to an argumentative essay. The second writer’s paragraph, on the other hand, begins by introducing the argument and briefly stating the beliefs postulated by both sides. She then indirectly reveals her own standpoint through reviewing her reasons for supporting one of the two sides.

5- Discussion

Writing has perhaps been one of the most difficult skills to teach in the foreign language classroom especially when academic preparation is questioned. Some view writing as a support system for teaching grammar and vocabulary, while others believe that writing deserves to be treated as an independent skill (Harmer, 2004). However, if writing were to be seen as an independent skill such as reading, listening and speaking, one would have to pinpoint the classroom practices and measures leading to the strengthening of this skill within learners.

Many textbooks adopting the skill-based approach tend to focus on teaching meta-linguistic strategies for writing (Byrne, 1988; Hedge, 2000; Hess, 2001). These strategies include the mechanics of writing, spelling, punctuation, generating ideas, and paragraph development. It appears that such an approach teaches learners how to organize their writing process but fails to sufficiently address the writing process itself.

The present paper introduces an approach to teaching writing which also considers writing as a stand-alone skill. However, it attends more to the linguistic
variables involved in teaching rather than the meta-linguistic ones which are currently in vogue. This approach makes extensive use of formulaic expressions and language chunks. The results of this paper revealed that such a formulaic approach results in greater coherence, grammatical variation and fulfillment of purpose by learners and accordingly greater academic accomplishments. Nonetheless, there was no significant difference between essays written by learners taught using this method and those who were taught through structures and grammar. Given the great importance ascribed to communicative competence in modern language proficiency tests, the authors of this article propose the use of the formulaic approach for teaching writing courses particularly in preparation programs of writing proficiency tests.

References


Mechanical system of life against humanity in Pirandello’s
I quaderni di Serafino Gubbio operatore (1925)1 (The
Notebooks of Serafino Gubbio Operator of Cinematograph)

Mohammad Hosein Ramazankiaei*
Assistant Professor
Faculty of Foreign Languages and Literature, University of Tehran, I.R. Iran
(Received: 8 Sep. 2008, Accepted: 3 Jan. 2010)

Abstract
This study investigates the process of dehumanization caused by a bizarre relationship between man and the machine. In the area of Italian literature there are very few novels that take into consideration the heady world of industrialization. Luigi Pirandello (1867-1936) one of the most famous European playwrights, novelist and short story writer places accurate emphasis on this problem. He chooses motion picture as symbol of mechanical system of life against humanity. The author insists that in the modern world the control we exercise over the work process through machinery undergoes a radical transformation. Mechanism becomes the principal means of controlling production and, as such, it controls human creativity. After the machine takes control of human creativity, man is no longer an individual and has no identity apart from the machine. The issue of industrialization against humanity is partially reflected in Pirandello's other novel The Late Mattia Pascal but it creates a bitter polemic in the author's sixth novel The Notebooks of Serafino Gubbio Operator of Cinematograph. This research focuses on the theme of life undermined by the voracity of industrialization. I additionally explained in the introduction that there is no affinity between Pirandello's theory of mechanical life and ideology of two contemporary thinkers of Postmodernism: Jean Baudrillard (1929-2007) and Paul Virilio (born in 1932).

Key Words: Pirandello, industrialization, Shoot, mechanical system, machine-made life, humanity.

1- From the film studio Kosmograph, Serafino Gubbio steadily winds the crank of his camera by day and scribbles with his pen by night revealing the world both mundane and melodramatic that unfolds in front of his camera. With his notebooks Serafino Gubbio parodies human weakness drawing attention to the theme of isolation and madness as emerging tendencies in the modern world. In various scenes of the novel Serafino mentions antithetical themes to show the impassiveness of the individual in modern society devoured by the machine. The machine makes human being so impassive that he gradually looses his homogeneity with other individual. In this case and as a result man finds difficulty to establish a regular relationship with others. Serafino confesses that he does not like to speak at all. He is aware of his difference from other individuals and knows that no one understands him because of his impassiveness. The main difference that Serafino finds between himself and other individual is that he finds others as persons but believes that his presentation resembles as that of a thing, a crank.
* Tel: 61119039, Fax: 021-88634500, E-mail:mkiiaeii@ut.ac.ir
Introduction

*The Notebooks of Serafino Gubbio (Shoot!)* was originally published in Italian in 1915. This novel is one of the first European novels to take as its subject the heady world of early motion pictures. The novel is a classic example of the Nobel Prize-winning Sicilian playwright Luigi Pirandello’s (1867-1936) literary talent for blurring the line between art and reality. From the film studio Kosmograph, the author’s character Serafino Gubbio steadily winds the crank of his camera and shows no tendency to change or fluctuate. He reveals a world both mundane and melodramatic that unfolds in front of his camera. Through Gubbio’s journal-like diary which is also saturated with fantasy, Pirandello struggles with the philosophical implications of modern industrialization. Like much of Pirandello’s work – such as *The Late Mattia Pascal* (1904) - *The Notebooks of Serafino Gubbio* reveals human weaknesses against the supremacy of the mechanical system of life. This is the machine that devours human life and makes man one of its innumerable parts. As far as the methodology of this research is concerned, I have focused on Pirandello's novel, in order to extract the antithetical matter of mechanical system of life, which is concealed beyond other issues the author is dealing with. Outwardly, Pirandello's theory of contrast between the machine and humanity is somewhat related to the French theorist jean Baudrillard and also to the French urbanist and philosopher, Paul Virilio. Baudrillard is a thinker of postmodernism who tried to combine philosophy and social theory in an original way while Looked at philosophy with contempt. In a letter to his son, Pirandello asked him not to waste his time studying philosophy because he has studied it and it has nothing to say. On the other hand, in his first full-length novel *L'Esclusa (The Outcast)* (1901) Pirandello again shows his disdain for philosophy. Baudrillard and Virilio focus on the postmodern and high technical society and find the distinction of race and gender as its structure. But Pirandello belongs to pre-modern society whose looking at contemporary society is merely the process of comparing the past and the present. Indeed Pirandello's feeling is nostalgic and he looks back nostalgically to what has been lost and does not pay attention to what is changing. Even Paul virilio as a cruel
critic against the postmodern industrialized and positive society does not feel nostalgia for the past. According to what Nicholas Zurbrugg says Virilio defines his general approach to contemporary culture (Zurbrugg, 193). Both Baudrillard and Virilio criticize disaster the technical progress brings for contemporary society. Even Baudrillard argues about the distance between an experienced truth and an obscured truth that the mass media claim to offer us (Tim Dant, 83). Here we encounter two different prospectives from two different group of thinkers. In the case of Baudrillard and Virilio, what deserves to be considered is cruel criticism versus institutionalised structure of contemporary society. But in Pirandello’s case we find just a feeling, a mixed feeling, of happiness, sadness and longing especially when it recalls a place or event from the past.

**Discussion:**

Pirandello’s disdain against rationalism has been embodied by description of a bitter inadequacy between man’s feeling and cruel mechanical life. Through his rational power man industrialises society in which his feeling is the first victim. The author’s reflection helps him to reveal not only the monstrous aspect of modern society but also facts that kills man’s spirit. The subject of industrialization and this doubt that whether it is able to guarantee a comfortable life is partially mentioned in the author’s Nobel Prize winning novel *Il Fu Mattia Pascal* (*The Late Mattia Pascal*)

1- When Mattia Pascal arrives in Milan he encounters a new atmosphere that completely differs from his simple homeland. Milan is a big and huge city and from the apparent aspect of that city Mattia Pascal deduces that all activities and generally life are under the control of a mechanical system. Mattia asked himself: “Perché’ tutto questo stordimento di machine? E che fara’ l’uomo quando le machine faranno tutto? Si accorgera’ allora che il cosi’ detto progresso non ha nulla a che fare con la felicita’? Di tutte le invenzioni, con cui la scienza crede onestamente d’arrichire l’umanita’ […] che gioia in fondo proviamo noi, anche ammirandole?” (Tutti i romanzi, Milano: Mondadori, 1966. p.353) Mattia Pascal has been deeply obsessed by this question: what man will do if machine does everything and takes man’s role in controlling life? And Pirandello’s reflection pushes him to ask: what man will do if machines somehow no longer are able to order society and life? Because, at first, man’s life did not depend on mechanical system; after man through science submits the control of his life to machine then he cannot detach it from its dependence.
Quaderni di Serafino Gubbio operatore has an immediate start to present the contrast between natural life and machine-made world. From the first page of the work the reader encounters Serafino’s impassiveness. Pirandello starts very soon to show that due to a mechanical life his character differs from other human beings.

"There is a something more in everything. You do not wish or do not know how to see it. But the moment this something more gleams in the eyes of an idle person like myself, who has set himself to observe you, why, you become puzzled, disturbed or irritated. I too am acquainted with the external, that is to say the mechanical framework of the life which keeps us clamorously and dizzily occupied and gives us no rest." (Pirandello, 1116)

Serafino’s statuesque pushes our reflection so that it recalls a free man whose life is destroyed by a mechanical system of modernisation. This catastrophic transformation has been described by Serafino himself:

"Man who first of all, as a poet, deified his own feelings and worshipped them, now having flung aside every feeling, as an encumbrance not only useless but positively harmful, and having become clever and industrious, has set to work to fashion out of iron and steel his new deities, and has become a servant and a slave to them'. (Pirandello, 1112)

Here Serafino describes two contrary states of man: man with his free feelings in a free existence and man as an object who is deprived of a natural life. Man as a poet defying his sentiment decided to throw this virtue away as a useless and harmful obstructive thing. As wise and industrious, he built his divinities of iron and himself has become their slave. The character continues with an exclamatory sentence: “Viva the machine that mechanizes life!” How can the machine have a
long life? The machine is voracious and insatiable that carries on its life by devouring man’s heart and soul. Serafino ironically asks the reader if he still has some soul or mind because the machine needs it. In Serafino’s question the emphasis is placed on the term “still” that satirically implies mechanical position of man deprived of soul and heart. There should be also a significant relevance between the author’s negative comment on the industrial civilization and the historic frame of the novel. In 1915 Pirandello published his novel originally entitled Shoot (Si gira). This happened exactly at the time of Italy’s intervention in the First World War and both of Pirandello’s sons had been summoned to enter military service. On the other hand, Futurists and war were quite intimately linked. They glorified war as the world’s only hygiene and as the destructive gesture of freedom that could easily destroy everything in order to prepare a fresh society for industrial civilization. Destroying libraries, museums and academies of every kind were among their manifests. Pirandello presented his utilitarian cowardice Serafino Gubbio at a time when Italian Futurists praised love of adventure, danger, violence and the mechanized onslaught of the war. But futurists did not praise war for itself and war was not their final goal. They believed in war as a movement that could sweep pre-industrial society preparing it for industrial civilization. A society occupied by the overwhelming presence of arsenals and shipyards blazing with violent electric, greedy railway stations, factories with lines of their smoke, bridges that cross over the rivers, adventurous steamers, locomotives and flight of the planes in the wind were futurists’ fervent dreams. ¹ Pirandello apparently brings Serafino as an anti-futurist who, paradoxically, returns to the past time when man, as a poet, was used to defy his sentiment. While Futurists excitement expresses their great approval for

militarism and patriotism, Pirandello’s character, secluded in his individuality, wants to return to the age of poetry.

**Man's presentation as a thing in the industrialized world**

Machine makes human being so impassive that he gradually loses his homogeneity with other individuals. In this case and as a result, man finds it difficult to establish regular relationship with others. Serafino confesses that he does not like to speak at all. (Ibid, 1134). He is aware of his difference from other individuals and knows that no one understands him because of his impassiveness. The main difference that Serafino finds between himself and other individuals is that he finds others as individuals but believes that he himself is just an object. Wanda Strauven says that Serafino's identity could not be detached from the handle he turns (Strauven, 127). He is an object, not a person and his identity depends on the identity of the object. And Fabrizio Denunzio remarks that Serafino Gubbio identifies himself with an object: his machine. (Denunzio, 191) One may add that even a machine has an identity but Serafino, as a part of a machine, has no identity far from that one defined by the machine. He has just a function. In the second section of the fourth notebook Serafino makes a comparison with his presence with Luisetta and three other persons in an automobile:

"She began to feel that my person was not necessary; but that my presence there had the necessity of a thing, which she as yet did not understand; and that I remained silent for that reason." (Ibid, 1178)

Serafino states that the young Luisetta at first knows him as one of those actors in the automobile who were speaking to each other. Afterward she realises that Serafino has no presence as a persona but as a thing. In fact Serafino’s utility as thing is necessary and not his presence as a man whose qualities can relate him as a

---

1- Ora cominciava ad avvertire che per questi altri ed anche per lei (in confuso) non ero propriamente uno. Cominciava ad avvertire, che la mia persona non era necessaria; ma che la mia presenza li aveva la necessità d'una cosa, ch'ella ancora non comprendeva; e che stavo così muto per questo.
human being. Stella M. John says that Serafino depicts himself as someone apart from society and estranged from others. (M. John, 52) Those present in the automobile can speak because they are persons but Serafino with his necessity of being a thing should be mute. This is the frustration of a man who subjugated his humanity to the machine and has been reduced to a mute object. Here Pirandello emphasises the alienation of man. The automobile can be considered as a metaphor in which an alienated camera man cannot be in touch with other individuals. The responsibility of alienation is attributed to the mechanised society as the source of alienation.

Once in The late Mattia Pascal, Pirandello condemned the machinery of being unable to generate happiness. In The Notebooks of Serafino Gubbio the Operator of Cinematograph the author deals again with antithesis between the machine and human nature. All actors of the Kosmograph studio travel with their own motor cars. But Serafino does not like to travel with motor cars which the film studio Kosmograph offers to actors. Travelling with cars does not let him see the beauty of trees on the roadside. He knows that instead of that beauty he will see shortly a monotonous and tedious green strip. Trees on the roadside which Serafino likes to see symbolises a peaceful and natural life that can provide enough happiness:

I, [...], am still here; I can console myself for my slow progress by admiring one by one, at my leisure, these great green plane trees by the roadside, not uprooted by the hurricane of your passage, but firmly planted in the ground, which turn towards me at every breath of wind in the gold of the sunlight between their dark boughs a cool patch of violet shadow: giants of the road, halted in file, ever so many of them, the open and uplift on muscular arms their huge palpitating wreaths of foliage to the sky(Pirandello, 1162).

And the tiresome strip of trees and other objects give a frightening sensation of
losing in speed. Pirandello focuses on the speed as a result of machinery in order to remind us of what we lose: nature. The silence of a natural and peaceful life has been broken by the sonorous noise of the engines of those motorcars the actors use for travel. In this novel Pirandello leads us to probe into the condition of human life. If in The Late Mattia Pascal a little canary imprisoned in a cage reminds the reader of the myth of freedom, here in Serafino’s diary book the reader finds all humanity totally entrapped in a huge cage. This cage of machinery not only embraces human being but also deprives him of all human senses. The problem is that the mechanism of modern life is based on those substantial elements that remain exactly on the opposite side of a simple life. The aspect of this mechanism will certainly eradicate those symbols through which man can refer to his nature. The row of beautiful shady trees changes into a humdrum strip. With this instance Pirandello invokes the reader’s mind to think about the noise of the machine and factories that completely veil the voice of birds and other elements of natural life.

The machine deprives human being of any human sense so that man becomes merely a contemplator and a simple observer who does not show any reaction in circumstances. According to man’s nature, in an emotional scene, in an exciting situation or even in a frightening circumstance, man’s various senses should be reactive. But when mechanization devours all these senses man’s reaction becomes mutinous. Ugo Olivieri reasonably argues that Serafino Gubbio loses all ties with his natural life in order to become a mute and alienated witness of a modern world dominated by technique. (Olivieri, 103) This is what Serafino testifies in that frightening scene where the beautiful actress Nestoroff and the actor Aldo Nuto die. Serafino stays in a cage displayed for the operator; he observes the scene without any reaction raised from human senses and captures mechanically the tragic scene by his camera. By describing Serafino in his cage, Pirandello exposes the deep contrast between man and his position in a mechanical system. In that cage Serafino lives out of time and also out of that common life with others. Without feeling and without any human senses the character’s brain has no function, he does not think and his heart never comes to be impressed by any emotional scene. He is a simple
crank and his eyes resemble a camera that does not lead to his heart whatever, though exciting, he encounters. Without any passion, compassion or ambition Serafino becomes a simple crank, a cold and impassive part of the machine. Douglas Radcliff-Umstead states that Serafino Gubbio is a dutiful servant of the life-killing machine. (Radcliff-Umstead, 251) Serafino’s instance explicitly shows that a dehumanised crank loses any possibility of being in communication with other individuals. If the relationship between the machine and man is essentially like that of an oppressor and a victim- or as the relationship between a devourer and a prey- the relationship between a victim and other people can be seen as that one of between a dead impassive individual with people who still have no conscience for dehumanisation. The machine withdraws those qualities of man that enable him to meet human needs or desires, those features that enhance his life. In this novel the impossibility of being in logical communication with other individuals does not derive from the eccentricity of life or multiplicity of personalities- an important issue focused on by Pirandello in his other novels such as *One, None and a Hundred Thousand*- but it has emanated from dehumanisation that deprives man of his individuality. Like a cold instrument man is devoid of any human senses and, as much as life and vitality is concerned he is dead. (Dombroski, 1994) In this world man is not eligible to fabricate his illusion and to build a fictitious world in order to put up with absurd life. He just serves machine as merely as a tiny part that in the next future will have not any successful function even for the machine. The vital force of self-creation that enables man to justify his existence comes to be denied by the machine. By presenting Serafino’s substantial transformation into a simple instrument Pirandello endeavours to testify verifiable facts that in the modern world of industrialization man is in danger being dehumanized by the machine. Unlike what the reader can encounter in Pirandello’s other novels such as *The Ltae Mattia Pascal* and *One, None and a Hundred Thousand*, in this novel the author makes use of a traditional literary style by presenting Serafino as a traditional narrative subject. Unlike Mattia Pascal of *The Ltae Mattia Pascal* and Vitangelo Moscarda of *One, None and a Hundred Thousand*, whose subjective position depends on the view
point of the other individuals, Serafino’s subjection can be nullified by insatiable and encompassing covetousness of the machine that has a strong desire to possess the control of human existence.

**An absurd coexistence between man and the machine**

Serafino’s attempt to write his diary is an attempt to make evident human feeling deformed by the voracity of the machine. In order to provide a reliable proof of dehumanization Serafino should objectivise himself as a crank. In fact a subject substance has been transformed into an object matter because of what the character wants to expose as the process of dehumanization. What the machine deforms with cruelty is not Serafino himself, but his subjective human substances such as feeling, love. In a likely metempsychosis Serafino’s human substances pass into the body of a crank.

The highest point of the novel appears in the final pages where the ever-contrasting coexistence between man and the machine emerge in the scene. The beautiful actress Varia Nestoroff and a tiger as symbols of naturality of life encounter the fierce nature of two rapacious machines: The gun and camera. Varia Nestoroff reminds the reader of man’s freedom entrapped by a film studio. She takes part in movies just in order to carry her life, and Anthony Caputi sees her as a victim of her helplessness. (Caputi, 251) She does not play role by any potential or instinctual force. Like Nestoroff, the tiger, too, is inflicted the compulsory presence in a film studio. Nestoroff is the prisoner of her mind undermined by the authorities of the Kosmografph movie studio and the tiger appears in a cage built by the machine. In this scene Pirandello shows how the insatiable voracity of the machine devours human life. The gun kills Nestoroff and tiger and Serafino the camera man captures this tragedy with his camera without any reaction. The prominent importance of this section exists in Serafino’s confession when he introduces himself as the slave of the camera:

"And I began to turn the handle […] My hand was impassively keeping the time that I had set for its movement, faster, slower, dead slow, as though my will had
flowed down--firm, lucid, inflexible--into my wrist, and from there had assumed entire control, leaving my brain free to think, my heart to feel; so that my hand continued to obey even when with a pang of terror I saw Nuti take his aim from the beast and slowly turn the muzzle of his rifle towards the spot where a moment earlier he had opened a loophole among the boughs, and fire, and the tiger immediately spring upon him and become merged with him, before my eyes, in a horrible writhing mass. Drowning the most deafening shouts that came from all the actors outside the cage as they ran instinctively towards the Nestoroff who had fallen at the shot, drowning the cries of Carlo Ferro, I heard there in the cage the deep growl of the beast and the horrible gasp of the man as he lay helpless in its fangs, in its claws, which were tearing his throat and chest; I heard, I heard, I kept on hearing above that growl, above that gasp, the continuous ticking of the machine, the handle of which my hand, alone, of its own accord, still kept on turning;"(Pirandello, 1274)1

Serafino's identity no longer differs from that of a camera and his existence appears as part of the camera’s existence and function. In fact Serafino’s human substance is amazingly incorporated into the mechanical function of a machine. For this reason Guido Baldi expresses that the main issue of the novel concentrates on the enslavement of man against the machine. (Baldi, 123) Serafino claims to be perfect not as a human being but as a component of a machine. Here Pirandello leads to a contrast between the machine and humanity to the highest point: the

1- "E io mi misi a girare la manovella,[…] La mia mano obbediva impensabile alla misura che io imponevo al movimento, più presto, più piano, pianissimo, come se la volontà mi fosse scesa - ferma, lucida, inflessibile - nel polso, e da qui governasse lei sola, lasciandomi libero il cervello di pensare, il cuore di sentire; così che seguitò la mano a obbedire anche quando con terrore io vidi il Nuti distrarre dalla belva la mira e volgere lentamente la punta del fucile là dove poc'anzi aveva aperto tra le frondi lo spiraglio, e sparare, e la tigre subito dopo lanciarsi su lui e con lui mescolarsi, sotto gli occhi miei, in un orribile groviglio. Più forti delle grida altrissime levate da tutti gli attori fuori della gabbia accorrenti instintivamente verso la Nestoroff caduta al colpo, più forti degli urli di Carlo Ferro, io udivo qua nella gabbia il sordo ruglio della belva e l'affanno orrendo dell'uomo che s'era abbandonato alle zanne, agli artigli di quella, che gli squarciavano la gola e il petto; udivo, udivo, seguitavo a udire su quel ruglio, su quell'affanno là, il ticchettio continuo della macchinetta, di cui la mia mano, sola, da sé, ancora, seguitava a girare la manovella."
machine at first deprives man of his humanity and then makes him become as a component of itself.

The diary of Serafino Gubbio not only depicts the process of dehumanization but also exhibits an encompassing moral void. Once the machine is made, it needs to devour man’s life and soul to act. Pirandello puts his stress on this importance that all our souls will be moulded in the same style so that if we put them on top of one another we can build a pyramid which can reach up to the stars. Serafino cannot avoid his anger, and pity as well, when he realises the actors of film studio submit themselves to the manufactured godhead of Kosmograph owing to their capricious ambition for wealth. This is a bitter trade in which man is caught. Based on the regulation of this treatment, man’s feeling comes to be considered as a useless thing that has no value. For this reason Serafino looks at people and at the airs they have and listen to them but he cannot believe in the reality of what he encounters. In *The Late Mattia Pascal* Pirandello wonders if the automation can provide happiness but here in this novel the author is convinced that modern civilization destroys human life. Pirandello affirms that a mechanism which can work itself without the interference of man, will be substituted by man. Man no longer will be of use and gradually the machine will eliminate him.

As mentioned above, since the first page of his novel, Pirandello places his emphasis on the machine as the source of man’s alienation in modern society. The protagonist proclaims against the mechanism of industrialization that restlessly occupies man’s life without offering any rest. The machine changes the quality of life and *fragore* (clamour) and *vertigine* (dizziness) (ibid,1109) become its inseparable features. These ambiguous features degrade the ingenuity of man in understanding the complicatedness of mechanised life. Serafino prefers to confess in the first page of his notebook the reasons for deciding to write down his experiences. In the contrast between humanity and the devouring machine, this is man that will certainly be defeated. His inability in defending his human qualities and senses against the devastating power of the machine brings about man’s madness. The character asks himself whether the clamorous machinery of life the day to day
becomes more complicated and whether it has not degraded man to a condition of madness which he must break out of and destroy everything (ibid, 110). Indeed Serafino’s notebooks present his outburst through which he not only makes reflections of his wretched plight but also of those victims who, like Serafino, are condemned to be nothing more than a silent hand in favour of the machine. I should agree with Roberto Alonge who states that the controversial polemic of the novel is not directed completely against the cinema (Alonge, 175) It may be truism to say that in Serafino’s notebooks cinema can somehow be the symbol of an artificial form of art. It exposes the character’s anxieties about daily industrialization seems to become more developed and complicated. Not cinema itself but cinematography as an industry and clamorous and dizzy mechanism of life brings about man’s folly.

Serafino’s nickname emanates from his role of automaton turning the crank without saying anything and without any eventual intervention. In fact the movie camera kills the life of two groups. At first Serafino who bends over the life-devourer machine becomes a mere silent hand with only a duty of turning the machine. And then the camera detaches the actors from their vitalities by showing them as void images on screen. Pirandello stresses that man in his industrialised life should submit all his innate powers to the power of the machine. In fact the machine has no faith in man’s creativity. When Serafino starts to operate the camera his eyes should follow the cold and indifferent eye of the machine. Man’s eye is able to both capture and comment on life but a camera’s eye only provides a dim and provisional draw of a scene devoid of any emotion. But this is man who should suppress his emotion in order to become an impassive thing. The camera imposes its mechanical function on man’s emotion, feeling and senses. After arriving at the state of a passive handle Serafino begins to record life. And his records are reflected in his notebooks. Here the character’s act resembles the function of a movie camera. A movie camera detaches life from itself as the machine records and finally reflects life deprived of its vitality on the screen. Serafino, too, records life and expresses it in the fragmented pages of his notebooks. Both the movie camera and the character
act as an indicator that shows impassively a detached and abstracted image of life.

A glance at Pirandello’s novels shows that the author believes in the alienation of the character in case of being absent from society. In fact the temporary non-existence of man results in the loss of those qualities which relate him to a particular place or atmosphere. In a psychoanalytical investigation made in the sixth book of Serafino the author emphasises a new feature in presenting this issue. After living with the memory of his grandfather’s house, it happens that Serafino visits the villa. At first he finds everything extremely changed but then he admits that the reality of the place has not changed and this is Serafino who has undergone by bestowing his humanity to the machine. The confusing and depressing fact is that the character does not know exactly whether he is changed or things have changed. The character makes a vain endeavour to create a sincere relationship with all those things which form his memorial life. Even a trivial memory of Serafino’s sweet past failed to relive or to revive that lost period because of his change:

"I had rung the bell. But sound was completely different. I could not realize that I was changed or sound of the bell was changed. How sad it was!" (Pirandello, 1235)

In his conversation with the tigress that the Kosmograph Studio bought to be used in a film, Serafino questions the culpable fierceness of man that with making use of the machine as instrument kills the innocent fierceness of nature. The animal’s fierceness derives from his nature but man’s fierceness emanates from his self-indulgence commented as self-defence:

“You cannot remain here on any other terms. Either you must be imprisoned like this, or you must be killed; because your ferocity—we quite understand—is innocent; nature has implanted it in you, and you, in employing it, are obeying nature and cannot feel any remorse. We cannot endure that you, after a gory feast, should be able to sleep calmly. Your very innocence makes us innocent of your death, when we inflict it in self-defence. We can kill you, and then, like you, sleep

1- "Sonai il campanello. Un altro suono. Ma ormai non sapevo più se dipendesse da me o perché il campanello era un altro. Che tristezza!"
Mechanical system of life against ... 163

calmly. But out there, in the savage lands, where you do not allow any stranger to
pass; not here, not here, where you have not come of your own accord, for your own
pleasure. The beautiful, ingenuous innocence of your ferocity makes the iniquity of
ours seem disgusting here. We seek to defend ourselves against you, after bringing
you here, for our pleasure, and we keep you in prison: this is no longer your kind of
ferocity; it is a treacherous ferocity! But we know, you may be sure, we know how
to go even farther, to do better still: we shall kill you for amusement, stupidly". (ibid, 1254)¹

Human beings destroy nature just for the simple reason of fun. The
impassiveness of man brings him to watch the destruction of nature and there is not
any connotation that can imply his pity or remorse. In fact this is man that by
submitting himself to a machine’s determination endeavors to exercise the machines
destructive power on nature. Thus what is more salient is man’s impassibility in his
cooperation with the tyranny of machine. Serafino’s stressis shown on some words
like “death” “kill” and he emphasizes that the tigress’s death enables the “pretty
machine” to extract profit. Pirandello shows the narrator as an ill character who
enjoys his role of being an impassive servant:

"You stare. At what do you stare, you beautiful, innocent creature! That is just
how things stand. You are here for no other purpose. And I who love and admire
you, when they kill you, shall be _impassively turning the handle of this pretty
machine here, do you see? They have invented it. It has to act; it has to eat. It eats
everything, whatever stupidity they may set before it. It will eat you too; it eats
everything, I tell you! And I am its servant. I shall come and plant it closer to you,

¹- "Tu qua non puoi stare altrimenti. O così imprigionata, o bisogna che tu sia uccisa; perché la tua
ferocia - lo intendiamo - è innocente: la natura l'ha messa in te, e tu, adoprandola, ubbidisci a lei e
non puoi avere rimorsi. Noi non possiamo tollerare che tu, dopo un pasto sanguinoso, possa dormir
tranquillamente. La tua stessa innocenza fa innocenti noi della tua uccisione, quand'è per nostra
difesa. Possiamo uccideriti, e poi, come te, dormir tranquillamente. Ma là, nelle terre selvagge, ove
tu non ammetti che altri passi; non qua, non qua ove tu non sia venuta da te, per tuo piacere. La
bella innocenza ingenua della tua ferocia rende qua nauseosa l'iniquità della nostra. Vogliamo
difenderci da te, dopo averti portata qua, per nostro piacere, e ti teniamo in prigione: questa non è
più la tua ferocia; quest'è ferocia perfida!"
when you, mortally wounded, are writhing in your last agony. Ah, do not fear, it will extract the utmost penny of profit from your death! It does not have the luck to taste such a dinner every day. You can have that consolation. And, if you like, another as well". (ibid, 1254)¹

The apogee of the character’s impassibility has been embodied by his insisting in getting closer to the tigress when she will be mortally wounded in order to plant his machine. Serafino’s grating depiction of the machine’s pitilessness and fierceness along with man’s passiveness continues when he ironically and satirically expresses that the tiger-skin will be presumably for Nestoroff the actress. She can use it as a costly rug under her little feet.

The silence of things which Serafino finally attains shows the protagonist’s endeavour to escape from the supremacy of the machine. Here I should refer to Ganeri's comment on this Pirandellian character. Ganeri sees Serafino as a little angel in a very small land who is on the exact opposite side of a huge industrialization. (Ganeri, 152) He is looking for a quiet home in sunlight, city empty of machine at night and a countryside far from the city where the noise of automobiles cannot break the silence. But finally he decides to have a mute presence. He prefers to speak to no one and his colleagues in the Kosmograph studio do not speak to him. They only convey to the protagonist those orders he can do without the necessity of giving any explanation. Douglass Radcliff-Umstead denotes that silence is a sentence that condemns Serafino. (Radcliff-Umstead, 267) The critic adds that the supremacy of the machine makes the character go berserk and writing notebooks is the result of his angry outcry. One may add that Serafino’s being completely a mute presence can be considered as his reaction against the

development of the machine as well as his attempt to compose the notebooks shows his outburst. With his silence Serafino can become a perfect automaton who should just turn the handle of camera. In fact silence indicates a process in which man undergoes a deep transition and passes from human being to the automaton. In this process man loses his human quality to become an impassive piece suitable for the machine. Indeed the protagonist’s silence is a way he chooses in order to put up with the agony of living. On the other hand it is the silence that enables Serafino to contemplate impassively the grating scene of man’s mutilation caused by the machine. Vila Fortuneti rightly says that Serafino Gubbio shows a dissociated and divided inner self since he passively witnesses on one hand, and criticizes society on the other hand. (Fortuneti, 277).

Conclusion

According to what Serafino Gubbio discusses in his diary, "life" and the "machine" are two different words that usually are referred to with two different meanings. Life refers to the natural origin of man and the machine concerns an industry that, with its insatiable voracity, devours man's life. There is something that, I would dare to say, the ambiguous conception that Serafino gives of life. As far as the machine is concerned, Serafino Gubbio gives a precise definition of machinery and shows explicitly how man becomes one of the trivial parts of the machine that make it work. But the novel does not focus on the meaning of some words such as "life" "humanity" and "creativity". What is the function of man's creativity that comes to be annulled by the machine? Is humanity a common quality of man's generation or are those individuals who avail themselves of humoristic vision to allowed criticize the modern world?

To sum up, the pathetic life of man starts when his life comes to be undermined by a mechanical system of life. Industrialization does not bring happiness for man but makes him become a component of the machine. In modern society and in life based on industrialization’ vitality of life is completely perished and existence depends merely on the function of machines.
References


A Cognitive Semantic Approach to Persian Spatial Prepositions, a Pedagogical Perspective, Case Study: Persian Preposition /dær/

Arsalan Golfam
Assistant professor, Tarbiat Modares University, Tehran, I. R. Iran

Fatemeh Yousefi Rad
Assistant professor, Payame Noor University, Qazvin, I. R. Iran

(Received: 1 July, 2008, Accepted: 15 May, 2009)

Abstract
We will argue that the common practice in the literature of giving a long list of senses for spatial prepositions, implying that they are unrelated to one another is not adequate. The more adequate way is presenting a network of related senses based on a prototype. Also, speaking specifically about Persian preposition /dær/, we will argue that the meanings of spatial prepositions arise as a consequence of our daily bodily experiences in the world and our conceptualization thereof. It is also shown that the cognitive semantic approach taken in this paper sheds more light on the semantic structure of spatial prepositions in general and /dær/ in particular, a light which can be very useful in teaching prepositions.

Key Words: Cognitive Semantics, Spatial Prepositions, Prototype Theory, Persian Preposition /Dær/, Language Teaching.
1. Introduction:

Persian spatial preposition /dær/ is roughly equivalent to English in and at. It is used so extensively that according to pldb\(^1\) corpus, it is the second most frequent word in Persian, preceded only by /va/ which means and. Also, to give more proof of the extended use of spatial prepositions, the third and fourth most frequent words in Persian according to pldb corpus, following /dær/ are /be/ and /az/, roughly equivalent to to and from respectively. According to Brala(2002, p.1):”Most, if not all (E)FL teachers and students are painfully aware of the fact that when it comes to mastering a foreign language one of the most troublesome areas to learn is the (idiomatic) usage of prepositions. Learning how to use prepositions correctly in a foreign language is a colossal task, one that is usually not an accomplished way of learning process, and one that many learners never manage to master thoroughly. As Lindstromberg (2001: 80) has pointed out, less than 10% of upper-level EFL students can use and understand prepositions correctly.”\(^2\)

Now, the traditional semantic analyses of /dær/ and other prepositions found in the Persian grammar books as well as the Persian dictionaries’ entries usually present a long list of senses, as if there is no relation between them. These accounts are a good source for the variety of usages these items have and also present a good historical review but tell us nothing of the high rate of systematicity found in the semantic structure of spatial prepositions in general and /dær/ in particular.

2. Theoretical Preliminaries

2-1) background

There are two main approaches to lexical semantics in general, classical and cognitive. On the whole, these two approaches differ in three main respects:

---

1- http://pldb.ihcs.ac.ir

a) Autonomy of linguistic semantics

Classical lexical semantics believes in the separation of linguistic semantics from encyclopedic information usually studied in pragmatics. So in this view, word meaning is autonomous. Cognitive semanticists, in contrast, suggest that all conceptual information associated with a lexical item is broadly encyclopedic in nature in that it is part of and needs to be understood against the background of broadly cognitive structures. They advocate that word meaning is not determined by the language system alone and that there is no clear line between purely semantic information and encyclopedic information.

b) Criteria for Defining Word Meaning: Necessary and Sufficient Features or Family Resemblance?

Following Aristotle’s definition of a category, classical lexical semantics considers necessary and sufficient features to be the requirement for category membership and hence, for the definition of polysemous items like prepositions. Cognitive semanticists in contrast, believe in Wittgenstein’s idea of family resemblance to be the motivation behind category membership:

“Indeed, members of a category in the extension of a lexical item may be linked, not because they all share the same criterial set of attributes, but because they share different sets of attributes with each other; in other words, because they are similar to each other in different respects”.1

“Similarly for spatial prepositions, it would be very hard to come up with a semantic description in terms of necessary and sufficient features. Indeed, the various usages or readings of a preposition are linked through similarity rather than through identity. The conceptual/semantic structure of spatial prepositions, then, can most appropriately be described in terms of a family resemblance network”.2

c) Prototype Theory

Cognitive semantics, unlike classical lexical semantics, believes in prototype

---

2- Hubert Cuyckens, ibid, p. 184
theory of categories. Put simply, prototype theory, originally put forth by Eleanor Rosch, maintains that members of a category do not all enjoy equal status. Some members are more central and some are more peripheral. So, membership is not a matter of one or zero but gradience is involved in the definition of category membership. Likewise, among the senses of a polysemous item like prepositions some senses are more central while others are more peripheral and link to the central sense through cognitive mechanisms like metaphor, metonymy, etc. Now, prototype categories have various characteristics which need not coincide. These characteristics include absence of necessary and sufficient definition, categorization based on salient member of the category, vague boundaries and radiality.

Put all these together, in cognitive semantic approach, the different senses of a polysemous item like prepositions are considered to form a family resemblance network. This is a prototype-based network the relations among its members are not arbitrary but highly motivated.

In this paper, this framework is adopted for the semantic analysis of Persian spatial preposition./dar/. The data is gathered mainly from contemporary Persian. Care has been taken not to base the discussions and conclusions on introspection. The reason is that:

“Cognitive linguists have strongly criticized the overwhelming use of introspection in linguistic methodology (Geeraerts, Grondelaers & Bakema 1994; Sandra and Rice 1995). The fact that theorists turn to introspective knowledge of the phenomena studied may lead to ad hoc conclusions. The two alternatives to introspection currently applied in cognitive linguistics are psycholinguistic experimentation and corpus analysis.”

Now, the method adopted here is that of corpus analysis.

2-2) Spatial Prepositions and Their Semantics

“Semantically, a preposition expresses a relation between two arguments x and...
\( y \), with \( y \) corresponding to that part of the prepositional constituent the preposition combines with and \( x \) being made up of one or more elements from the rest of the sentence or the neighboring discourse that contains the head of the constituent the PP is a complement or an adjunct to. Spatial prepositions, now, indicate the spatial relation between two arguments \( x \) and \( y \), i.e., how \( x \) and \( y \) relate to each other in space. More specifically, in a large number of cases, they describe the place/location of \( x \) (in other words, they assign \( x \) to a particular place) by using the argument \( y \) as reference, or still, they serve to locate \( x \) with respect to \( y \), or rather the place of \( y^{1} \). Now, in cognitive linguistics, the terms trajector (TR), and landmark (LM) are used to refer to \( x \) and \( y \) respectively.

Before the advent of cognitive semantics, the meaning of words in general and spatial prepositions in particular, was defined on a necessary and sufficient criteria basis (Cooper, 1968, Jackendoff, 1976, Kats, 1966 & 1972, Leech, 1969, Tarsky, 1956).

According to this tradition, a unique set of necessary and sufficient criteria accounts for the meanings of each word. However, highly polysemous words, and most notably among them prepositions, count as a serious challenge to this explanation. How should one define a set of necessary and sufficient criteria so that it can cover all and only the different senses of a highly polysemous spatial preposition?

Now, within the framework of cognitive semantics, prepositions receive a much better semantic explanation. Cognitive semantics proposes that different senses of a polysemous word are linked through a highly-structured prototype-based semantic network. This paper, in the spirit of Cognitive Semantics, claims that the different meanings of Persian preposition /dær/ form a radial network, the links among it’s members are not arbitrary. Quoting Brugman and Lakoff(2003):

“The theoretical claim being made is that a polysemous lexical item is a radial category of senses. What is important for our purpose is that the kind of network

\footnote{Hubert Cuyckens, “The Semantics of Spatial Prepositions in Dutch”, 1991, pp. 80-81}
structure found here is not made up ad hoc to characterize this set of facts. Instead, this is a common category structure that occurs in domains other than the lexicon. There is an important consequence of using the general theory of radial categories to characterize polysemy. In the general theory, the links between members of the network are not arbitrary. The theory of radial categories comes with a characterization of possible link types. In the case of polysemy, the link types are the types of relations linking the senses of the word. In general, some of the links may involve shared information, some may involve relation between a general and a specific case, and some may be metaphorical…. But, overall, there is only a small number of types of relations between senses of words….”

3. Methodology

3-1) First, the sentences and phrases containing /dær/ were gathered. The data were primarily collected from contemporary Persian utterances. Also we tried to abandon reliance on introspection as far as possible.

3-2) After collecting the data, the many usages of the preposition were categorized according to factors such as LM configuration, the relation between LM and TR, both geometrically and functionally. The same was done also for temporal and abstract usages of the preposition in question beside the spatial usages.

3-3) Then the prototype sense of /dær/ was identified. This prototype sense is presumably the spatial sense. However, care should be taken that even among the strictly spatial senses of the preposition, there is a considerable variety of usages and singling out the prototype sense is not very straightforward. So, even among the spatial senses, one can observe centre-periphery structure, some usages being more central than others.

It should be noted that the selection of spatial meaning as the prototype sense is not arbitrary, having its root in localist hypothesis. However, there are some other independent proofs regarding the prototypicality of spatial sense, among these one

can mention the frequency analysis of corpus and the psycholinguistic experiments, not the subject of this study.

3-4) Then, the relation among the senses was characterized. We believe that only a limited number of relations are possible among the different senses of a polysemous item and these relations are the same processes playing role in diachronic semantic change, esp metaphor, metonymy and also semantic extension and narrowing.

3-4) Finally, taking into consideration the existing links among the different senses, the semantic network of the preposition was proposed.

Concerning the limitations of this study, it should be noted that although prepositions including daer are actively involved in the creation of phrasal verbs, these phrasal verbs are not studied here. Moreover, complex prepositions including daer are also excluded from the scope of this study. This by no means implies their theoretical triviality, but on the other hand they deserve a separate study in their own right.

**Data Analysis**

In presenting the data, we have chosen to translate the phrases and sentences containing /dær/ but to put /dær/ intact. As hinted before, it is usually translated as in or at.

Now consider some of the data:

Water is daer the pot.
He is daer the office.
I am daer the street.
The president daer the head of a high ranking delegate went to the UN.
The three Finnish militia daer Iran waters
The bomb which was planted daer the route of the car
He is daer my hand/ fist (metaphorically, completely under my control)
Daer times like….
Daer the age of 15
The petroleum station burning der fire

this moment

that point

A short look at these examples of the usage of /der/ is a sign of its wide application. Der is used not only with spatial, temporal and abstract LMs, but within each of these domains exhibits a considerable variety. To give an example of spatial domain, the kind of relation the LM of der has with its TR is highly flexible: pot is a 3-dimensional entity/container which contains the TR (water) while Iran waters is a rather shapeless liquid. Street is a 2-dimensional surface and point is well, just a 0-dimensional point. This variety is also exhibited in temporal and somehow less in abstract domains.

Now the question is, how to define one meaning of /der/ that can cover all this variation. In other words, taking into account that usually the LM features are considered as part of the meaning of prepositions, what kind of meaning can be posited that includes all and only the above configurations?

It seems that we should abandon the endeavour to find a unique definition based on necessary and sufficient criteria for der and we should consider /der/ as having a semantic network with multiple nodes, the links among which are established through family resemblance, that is to say, the adjacent nodes are the more similar meanings of the preposition, having some common features while the farer nodes not necessarily having common features.

Now, the question is how to account for the links between adjacent nodes in the polysemous network of /der/?

This paper tries to show that the different senses of /der/ form a radial network the relations between whose nodes are not arbitrary but are principled and recurrent through the lexicon. So they form a highly structured system with the links between its nodes being motivated. Let’s see how.

After investigating the data, a pattern of usages shows itself. Within the domain of spatial usages, der is used with 3-, 2-, 1- and 0-dimensional LMs. As an example
A Cognitive Semantic Approach to ...

of each usage, one can mention respectively:

a. Water /deer/ the pot
b. /deer/ the TV screen
c. The bomb planted /deer/ the route of … car
d. Yazd is /deer/ the centre of Iran

Now, one may ask why we should consider route in c as an example of 1-dimensional LM since in real world no route is actually one dimensional, routes and streets usually having width as well as length. The answer lies in one of the main tenets of cognitive semantics, ie the process of conceptualization. Hence, in the cognitive semantics approach to language, it is believed that language is not a mere mirror to the world out there, but what is more important is the way we as humans conceptualize it. Now in the process of conceptualization, depending on the situation one may decide to make some elements of reality more prominent and push some other presumably less important features to background. So in the example above it is the one dimensionality and linear feature of route that is at issue and hence brought into prominence. The same explanation goes for d. The city of Yazd has not presumably occupied just one point on the map of Iran, however in this context it is just its location that is at issue and hence it is conceptualized as a 0-dimensional point.

Now, let us consider /deer/ in its temporal usages. It is interesting to note that in its temporal usages, /deer/ has a more or less similar configuration to that of the spatial ones and this is encoded lexicalized very nicely and clearly in Persian:

1. /deer/ the container of 3 hours
2. deer the length of last 5 years (referring to the duration of time)
3. deer the width of 5 minutes (referring to time limit)
4. deer that point of time/ moment

So, in the first example, the domain of time is conceptualized as a 3-dimensional entity and this is clearly lexicalized through the word container. In the second and third example, time is conceptualized as a line, hence one-dimensional and finally in the fourth example time is conceptualized as a 0-dimensional point.
It is clear that one of the distinctions which was at work in the spatial domain, i.e., the distinction between 3- and 2-dimensional LMs disappears in the domain of time, hence the preposition *dar* conceptualizing time into a 3/2 dimensional, 1-dimensional and 0-dimensional.

But what about the abstract domains? What does *dar* refer to in the more abstract cases of usage? Let’s see some examples:

* /dar*/ the domain of politics
  She has no competitor /dar*/ intelligence
  /dar*/ the result of cancer, …
  /dar*/ situations which…

It is clear that in the sphere of abstractions, the classification of usages/tokens dimensionally becomes more difficult, even impossible. What is important here is not geometrical features, but rather the functional features of the relation introduced by /dar*/. A central relation at work here is that of *control*. In this relation, usually LM, controls the location of TR, in the same way that a container controls the location of the contained material. Moreover, in explaining the abstract cases one can employ the conceptual metaphors introduced by Lakoff and Johnson, specifically the conceptual metaphor *states are locations* accounts for the cases in which /dar*/ is applied for states of politics (a) and being intelligent (b).

Now, what has been said so far can be summarized in the following table:

<table>
<thead>
<tr>
<th></th>
<th>Spatial domain</th>
<th>Temporal domain</th>
<th>Abstract domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>3- DIM</td>
<td>inclusion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2- DIM</td>
<td>boundedness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1- DIM</td>
<td>Position/ place(linear conceptualization)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0- DIM/ point-like</td>
<td>Position/ place (punctual conceptualization)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Notes:

1. It is clear that in the domain of time, the difference between the 3-dim LMs and 2-DIM LMs is lost.

2. The bold arrow at the left side of the table indicates the level of control which LM exerts on TR. Note that the more we move away from 3-DIM LMs which act as containers for their TRs, the less the amount of control exerted from LM into its TR.

Another important point to be noted is that within the theoretical framework adopted in this paper, the question of how many senses the preposition /dær/ has is not a theoretically important one for several reasons. For one thing, this question is based on the conduit metaphor of language, according to which words act as containers for meanings and meanings are regarded as things contained in words. Now, this view is not maintained in cognitive semantics. In the view of cognitive semantics, meanings are not regarded as stable entities inside words but are contextually flexible. For another thing the answer to the above question varies according to the level of abstraction one decides to adopt for their semantic analysis, the more abstract and schematic the level of analysis, the less the number of meanings would be.

4-1) The Prototype of /dær/

The prototype sense of /dær/ is presumably its spatial sense and among the several usages of deer in spatial domain, the one in which LM is 3-dimensional and plays the role of container for it’s TR is the prototype sense of /dær/. In this prototype sense, the functional element of control goes hand in hand with the geometrical element of 3-dimensionality and inclusion. There are some clues in determining the prototype sense of polysemous words in general, one is higher frequency of usage of the prototype as compared with other senses, another the native speakers intuition. This second clue demands some explanation especially in the light of what we have said before concerning the attitude of cognitive semantics towards introspection.

According to Geeraerts (2006), given the presupposition that introspection
yields only a partial insight into the semantic structure of words, we can also expect that it is exactly the prototypical kinds of usage of those words that reach the introspective consciousness of language user. So as far as the determination of prototype is concerned, we can also partly rely on intuition. The intuition also clearly points to this usage as the prototype of *dær*.

Moreover, the spatial senses of *dær* are motivated through *time as space metaphor*. So, spatial senses are extended by the use of this conceptual metaphor to yield temporal senses.

Now, it seems that the abstract senses of *dær* are extended from the geometrical and functional components of its prototypical spatial sense. So we propose the following model for the extension of abstract senses of *dær*:

<table>
<thead>
<tr>
<th>Components of spatial <em>dær</em></th>
<th>Abstract extensions</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boundedness ( through metaphor)</td>
<td>State/ situation sense</td>
<td><em>Dær war</em></td>
</tr>
<tr>
<td>Inclusion ( through narrowing)</td>
<td>Part- whole sense</td>
<td><em>I am dær the army</em></td>
</tr>
<tr>
<td>Physical control (through metaphor)</td>
<td>Metaphoric control</td>
<td><em>He is dær my fist!</em></td>
</tr>
<tr>
<td>Surrounding ( through metonymy)</td>
<td>cause</td>
<td><em>Dær the result of…</em></td>
</tr>
</tbody>
</table>

An important point to be noted here is that although we suggest that such- and-such a sense is taken from the component $x$ of prototypical *dær* does not imply that the other components of spatial *dær* are not present in that sense. Rather, the question is merely the relative salience of that component in such a specific sense.

5. Conclusion

The proposed radial network of the prototype-based family resemblance structure of /dær/ should not –and as we see it, cannot - be considered as a mere rephrasing of the traditional accounts of the meaning of this preposition because it actually does more than provide a mere list of contexts and their meanings. It also
explains the semantic contribution of /dɛr/ to those contexts. Moreover, it also provides a scheme that explains the conceptual links between different senses, thus rejecting the common—even if implicit—idea that the meanings are quite arbitrary.

Although it is not always straightforward to identify the links among the senses of a preposition, it is also true that in the majority of cases it is possible. Hence what appears to be completely arbitrary in traditional accounts turns out to have a considerable degree of structure and motivation in cognitive semantic framework. In this way, less burden is put on the memory of language learner and the semantic structure of preposition seems much more systematic, hence learnable.

References


Ignasi Navarro I Ferrando. (2002), "Towards a Description of the Meaning of AT", in *Perspectives on Prepositions*, Edited by Hubert Cuyckens and Gunter Radden, Tubingen.


"Imaginative Geography": Orientalist Discourse in Paradise Lost

Seyyed Mohammad Marandi*
Assistant Professor University of Tehran, I.R. Iran

Hossein Pirnajmuddin**
Assistant Professor English Department Faculty of Foreign Languages University of Isfahan, I.R. Iran

(Received: 29 Oct. 2008, Accepted: 13 Jan. 2010)

Abstract

Paradise Lost incorporates many references to the East. The Orient figures prominently in the vast scope – the "imaginative geography" - of the poem. This paper attempts a survey of what, following Edward Said, has been termed "orientalist discourse" in Milton's epic poem. It is argued that this discourse has to be considered in the context of Milton's essentially religious and anti-monarchical stance. Associating the Orient with evil and the Satanic regime Paradise Lost cannot be wrested from "latent orientalism" but it is shown that issues such as aesthetic considerations, a cosmic setting, drawing on the authority of history, classicism, an encyclopedic scope, an essential anti-monarchism and above all a profound process of displacement whereby comments on contemporary issues are displaced onto the Orient all help compound the representations of the East in this text. The result is an ambiguous and multi-faceted orientalist discourse.

Key Words: Milton, Paradise Lost, the Orient, Orientalist Discourse, Colonialist Discourse.
Introduction

Edward Said’s pivotal thesis in *Orientalism* is that texts of orientalism ‘can create not only knowledge but also the very reality that they appear to describe’ (1978, p. 94). Orientalism is above all ‘a style of thought,’ (ibid, p. 2) a way of thinking about and "imagining" the East. Elsewhere Said writes of the "imaginative geography of 'our land-their land' variety" which is a way of "designating in one's mind a familiar space which is 'ours' and an unfamiliar space beyond 'ours' which is 'theirs' ” via which human societies seem to "derive a sense of their identities negatively" (1978, p. 54). The orient figures quite noticeably in *Paradise lost* and *Paradise Regained*. This paper is concerned with Milton's "imaginative geography" of the Orient in *Paradise Lost*. In a survey of the representations of the East in *Paradise Lost* it is argued that Milton’s orientalist discourse, which in itself reveals some deep-seated conceptions about the East, serves to refract the author’s political views and sharpens the thematics and imagery of his poem. Of course, it is also noted that these images of the East are aesthetically functional too. As such, the notion of orientalism as a discourse of Western hegemony is compounded.

Although no comprehensive study of the representations of the East in Milton is done to date, a number of scholars have touched upon the issue. Martin Evans's *Milton's Imperial Epic, Paradise Lost and the Discourse of Colonialism* (1996) is probably the most comprehensive single study approaching the subject. But the focus, as the title indicates, is colonialist rather than orientalist discourse. Davies (1983), Zwicker (1987), Lim (1993), Quint (1993) and Stevens (1997) discuss the issue along similar lines. Oddly, for a discussion of representations of the East in Milton per se one has to refer back to Samuel Chew's *The Crescent and the Rose* (1937), a monumental work which traces the representations of the Islamic East in Renaissance Europe. The present article aims at discussing the significance of such representations in Milton's epic poem.

Discussion

One of the striking aspects of *Paradise Lost* and, to a lesser extent, *Paradise
Regained is their wealth of geographical detail, their truly vast ‘imaginative geography,’ as Said puts it, including almost all the known world as well as the ‘discovered’ lands of that time. The vast geographical perspectives of *Paradise Lost* embrace Asia, Africa, Europe and Americas (from "Cambalu…And Samarchand by Oxus…" to "Agra…Mosco…Mombaza, and Quiloa, and Melind…On Europe…Rich Mexico…El Dorado")\(^1\), lands imagined and real, charted and uncharted. The East of course figures prominently in this grandiose setting.

To begin with, why include Milton in the orientalist canon at all? What were the reasons for, and what purposes were served by, the inclusion of the matter the Orient in *Paradise Lost* (and its sequence, *Paradise Regained*)? Why does the Orient figure so significantly in what can be called the geographical or cartographical discourse of these poems?

It is no wonder to find Orientalia in a poem which, to use Barbara Lewalski’s words, has "an edenic profusion of thematic and structural elements" (199, p. 113). The orientalist discourse is arguably deep-seated, embedded in the structure and buried within the allusive texture of these poems. Figuring, however, in the work of, as Alistair Fowler (1998, p. 41) puts it, "one of the most politically engaged of all [English] poets," in texts which are, in Tony Davies’ phrase, "the battleground of competing ideologies," (1992, p. 22) this discourse has to be considered in negotiation with other discourses circulating at the time of the production of these texts. Here the New Historicist notion of the literary text is useful -- the idea that literary texts to some extent map the discourses circulating at the time they are written and, to some sense, are themselves part of these discourses. This figuration can be accounted for in a number of ways.

First of all, these images of the East are aesthetically of significance. Critics have noted the abundance of proper names, geographical and otherwise, in *Paradise Lost*. Fowler (1998, p. 21) discusses T. S. Eliot’s and F. R. Leavis’ treatment of the

"'roll-call' of names as having only a vague atmospheric or auditory value" and goes on to explain how the recent study of theological and historical contexts shows that "every naming…is sharply edged." But this idea of "atmospheric or auditory value," is, we think, something quite relevant to the discussion of orientalism in these poems. "Samarchand" (once a celebrated city in the Persian empire), "Cambalu," "Ecbatan" (PL, XI. 391) (an ancient capital of Persia), "Hispahan" (PL, XI. 392) (the capital of Persia in the Safavid era, at Milton’s time) are names, which, apart from other associations, are sonorous, evocative, even romantic. The note of wealth, exoticism and glory associated with the earthly kingdoms is certainly there, but one is conduced to assume that these oriental, along with the non-oriental ("Quiola,' ‘El Dorado," etc.), names to some extent are there to achieve atmospheric and sound effects befitting an epic on a cosmic scale.

Also, these texts claim to render universal themes, the story of man’s creation, his fall and his final redemption. These themes are presented on a cosmic scale. Furthermore, the poems treat biblical stories and the biblical lands are located in the East. For instance, "With these came they, who from the bordering flood / Of old Euphrates to the brook that parts / Egypt from Syrian ground…” (PL, I.420-23).

Milton was also writing in the late Renaissance, a time of cartographical awareness, a time of, in Stephen Greenblatt’s words, "heightened interest in boundaries" (1980, p. 11). It was a time of exploration, map-making and discovery. These boundaries, however, were mostly blurred and many parts of the globe were to the West an "as yet unavailable terra incognita" with "its people and customs often the product of European fable, fourth-hand report" (Davies, 1991, p. 123). The Renaissance discourses of exploration, discovery and cartography, then, impinged on Milton’s conception of geographical places as the, to use Stevie Davies’ words, "inconclusive” geography of Paradise Lost evidences (ibid.).

This geographical discourse, with its inclusion of diverse lands, many of them Eastern, is intertwined with the thematics of these poems. For instance, in Book XI of Paradise Lost there is a catalogue of the Eastern empires. Adam is shown
The seat
Of mightiest empire, from the destined walls
Of Cambalu, seat of Cathaian Khan
To Paquin of Sinaean kings and thence
To Agra and Lahor of great mogul
Down to the golden Chersones, or where
The Persian in Ecbatan sat, or since
In Hispahan, or where the Russian Czar
In Mosco, or the Sultan in Bizance.
Turkhestan born; (XI.386-96)

Fowler (1998, p. 23) observes that here there is a sort of "topomorphic" or "spatial" pattern. Such patterns, he adds, in *Paradise Lost* often imply "symbolism of sovereignty -- not of course to dignify human monarchy, but rather to affirm Christ’s kingship and the moral hierarchy of creation." Chersones, identified with Ophir, which supplied Solomon, the only biblical ruler mentioned in the poem, with gold, is centrally placed, flanked with "the Mogul’s and the Persian’s realms, each with two capitals; then the Sinaean kings’ and Czar’s, each with one capital" (Fowler, 1998, p. 618). The matter of the Orient is used to drive home the focal theme of the poem: the denunciation of worldly kingship and underscoring God’s sovereignty. The orientalism of the poem here is inseparable from its thematics, it helps give aesthetic form to it.

**Classical, biblical and contemporary allusions**

The matter of the orient incorporates images of exoticism, splendour and wealth, kingship and despotism and barbarism at a deeper level (in Stevie Davies’ words, images of "formless, chaotic masses of beings on the borders of the human, barbarians who pour down upon civilization…’ (1983, p. 93)). One way to approach these images and allusions, we propose, is to group them into three main frameworks (which at times overlap): classical, biblical and contemporary. These
allusions are part and parcel of the central theme of Milton's epic poem: the
denunciation of worldly kingship and the glorification of the "Kingdom of Heaven."
Here is, for instance, an image of oriental luxury and splendour used in describing
the Pandaemonium:

Not Babylon,
Nor great Alcairo, such magnificence
Equalled in all their glories, to enshrine
Belus or Serapis their gods, or seat
Their kings, when Egypt with Assyria strove
In wealth and luxury. (I.716-21)

The pagan East is depicted in its Satanic luxury. The above image is about an
ancient, pre-Christian era, the following is that of a more contemporary one. Satan’s
regiment throng when a council is proclaimed in hell:

They anon
With hundreds and with thousands trooping came
Attended: all access was thronged, the gates
And porches wide, but chief the spacious hall
(Though like a covered field, where champions bold
Wont ride in armed, and at the soldan’s chair
Defied the best of paynim chivalry
To mortal combat or career with lance) (PL, I.759-66)

"Soldan," resonating with associations of tyranny and cruelty, and "paynim" are
markers of orientalist discourse here. The image is specifically that of the Islamic
Orient and to be more exact that of the Turkish rulers. Elsewhere the Satanic council
is likened to the typical oriental state councils. Satan re-enters the Pandemonium:
"Their mighty chief returned: loud was the acclaim:/ Fortho rushed in haste the great
consulting peers, / Raised from their dark divan, and with like joy" (PL, X. 455-57).
"Divan" is a Persian word that, among other things, means a state council as well as,
ironically, a council of justice. The alliterative "dark divan," resonating with overtones of corruption continues the implications of the previous image.

The devils can also be likened to an obscure people somewhere in the East:

So thick the airy crowd
Swarmed and were straitened; till the signal given,
Behold a wonder! They but now who seemed
In bigness to surpass Earth’s giant sons
Now less than smallest dwarfs, in narrow room
Throng numberless, like that pygmean race
Beyond the Indian mountain, (PL, I.775-81)

The oriental motif continues and the "numberless" devils are likened to a barbarous Oriental people from beyond the mountains, "formless, chaotic beings." This blurred geography is invoked in a series of epic similes associating Oriental territories with the Satanic regime. For instance, the encounter between Satan and Death (his son by his daughter Sin) is described as the collision of two clouds over the Caspian in the north of Persia (PL, II.714-16). Or, Satan’s roaming of the earth is described as follows:

Here walked the fiend at large in spacious field,
As when a vulture on Imaus bred,
Whose snowy ridge the roving Tartar bounds,
Dislodging from the region scarce of prey
To gorge the flesh of lambs or yearling kids
On hills where flocks are fed, flies toward the springs
Of Ganges or Hydaspes, Indian streams; (PL, III.430-36)

However, oriental realms are not always associated with Satan and hell. Images of the exotic Orient are deployed to describe heaven as well; heaven’s balmy scents are likened to:
As when to them who sail
Beyond the Cape of Hope, and now are past
Mozambic, oft at sea north-east winds blow
Sabean odours from the spicy shore
Of Arabie the blest, with such delay
Well pleased they slack their course, and many a league
Cheered with the grateful smell old Ocean smiles. (PL, IV.159-65)

Several discourses overlap here. "Cape of Hope" and "Mozambic" as well as the whole of this nautical image are markers of a discourse of mercantilism and discovery dominant at that time. "Sabean odours" betokens the biblical discourse (Saba being a biblical land) and "Arabie" marks orientalist discourse. The oriental image of the passage helps create a sort of atmospheric effect, something related to the aesthetics of the text in general that will be discussed at some length later. Likewise is the description of the scene in which Eve picks the delicacies of the Heaven: "Whatever earth all-bearing mother yields / In India east or west, or middle shore / In Pontus or the Punic coast," (PL, V.338-40). The East is imaged as a locale of edenic profusion and abundance. One more image of this kind, again specifically of India, is used in the scene in which Adam and Eve cover themselves with fig leaves: "The fig tree, not that kind for fruit renowned, / But such as this day to Indians known / In Malabar or Decan spreads her arms" (PL, IX.1101-103). Fowler (1998, p. 502) notes that Purchas could be the possible source for this image. The deictic "this day" shows the wide circulation and direct stamp of the discourses of exploration and mercantilism, the Hakluyt and Purchas collections, on the Miltonic text. Such images indicate that much of the orientalism of the text under discussion is engaged with the aforesaid discourses rather than the discourse of colonialism.

To highlight the significance of such images we proceed to our proposed classification (classical, biblical and contemporary). An important instance of the classical framework is the allusion to Xerxes’ Greek expedition. Sin and Death
build a bridge over the chaos linking hell and earth:

   Over the foaming deep high arched, a bridge  
   Of length prodigious joining to the wall  
   Immovable of this now fenceless world  
   Forfeit to Death; from hence a passage broad,  
   Smooth, easy, inoffensive down to hell.  
   So, if great things to small may be compared,  
   Xerxes, the liberty of Greece to yoke,  
   From Susa his Memnonian palace high  
   Came to the sea, and over Hellespont  
   Bridging his way, Europe with Asia joined,  
   And scourged with many a stroke the indignant waves. (PL, X. 306-11)

Milton is here drawing on Aeschylus and Herodotus. Susa (the winter palace of the Persian kings known to the Greeks as Memnonia after Memnon, the son of the dawn, Aurora) is analogous to hell and Xerxes to Sin and Death, the offspring of Satan. On the aptness of this image, Fowler (1998, p. 557) observes that "Death and Xerxes both build bridges, intend to subdue nations, are proud and strike the deep." The passage reproduces two of the salient topoi of the classical Greek writers’ discourse on Persia, tyranny and pride. These two topoi are specifically highlighted. Xerxes builds his bridge "the liberty of Greece to yoke" and his overweening pride makes him scourge the sea in indignation. The reproduction and recirculation of these topoi empower and authorise the Miltonic text. Moreover, as Xerxes’ invasion of Greece, recorded in Herodotus, is regarded as a historical fact, the passage also draws on the discourse of history and hence doubly empowers the text. Paradise Lost, says Mary Radzinowicz (1987, p. 206), is "a course in political education." Milton tries to "educate" his readers by underlining the inseparable link between tyranny and kingship through the discourse of history. Milton’s use of the matter of the Orient, then, is related to his particular conception of history as his works, holds C. A. Patrides (1997, p. 269), are "the most successful attempt in poetry to fuse the
essential aspects of the Christian view of history into a magnificent whole…Above all we have the universalistic and Christocentric view of history." Milton suggests, and "educates" his readers as to the fact, that history shows that except the one true sovereign, Christ, all kings are fraudulent and tyrannical.

This classical-orientalist discourse is also related to what can be termed Milton’s display of erudition. In a seminal essay on the relationship between politics and the literary culture in the Restoration, Steven Zwicker (1987, p. 259) shows how Milton tried to give authority to his voice by drawing on "the language of cultural authority -- Scriptures, the classics, scholarship, history, literary learning." Milton, the staunch champion of the overthrown republic, surrounded by his Royalist enemies, was trying to assert his voice, "out of humiliation and political displacement, out of the need to assert the authority of his voice came the epic scope of this poem" (Zwicker, 1987, p. 248). The Orientalia of Paradise Lost (as well as those of its sequel, Paradise Regained), the bits and pieces of "knowledge" about the history, geography and customs of the East, come out of these languages of "cultural authority." As regards the aforementioned wealth of geographical detail in these works, Milton’s attempt at an encyclopaedic scope had to incorporate cartography as well as it was an important field of knowledge and erudition at that time. To return to the image under discussion, there is a resort both to the discourse of history (Herodotean account of the Greco-Persian wars) as well as that of the classics (Aeschylean account), which here happen to be more or less overlapping. Here, as in Paradise Regained, the orientalism of the text gives an edge to its pivotal theme, an indictment of all kingship except the kingdom of God. This brings us to the discussion of a cluster of images of kingship and empire that comprise many of the allusions to the Orient.

As critics have noted, kingship in Milton’s two major works is mostly associated with the satanic regime. For instance: "Satan, whom now transcendent glory raised / Above his fellows, with monarchical pride" (PL, II.426-28). Or: "Thus saying rose / The monarch [Satan]" (PL, II.466-67). Satan is also called the
devils’ “great emperor” (PL, I.379). Perhaps the best encapsulation of Milton’s view of the kings is the scene in which Adam in his vision denounces Nimrod:

O execrable son so to aspire
Above his brethren, to himself assuming
Authority usurped, from God not given:
He gave us only over beast, fish, fowl
Domination absolute’; that right we hold
By his donation; but man over men
He made not lord; such title to himself
Reserving, human left from human free (PL, XII.64-71)

The denunciation of kingship is mediated through biblical (the above example), classical (the image of Xerxes’ expedition, for instance) as well as contemporary frameworks. In the following lines Satan and his fellowship are imaged as a sultan and his host:

As when the potent rod
Of Amram’s son in Egypt’s evil day
Waved round the coast, up called a pitchy cloud
Of locusts, warping on the eastern wind,
That o’er the realm of the impious Pharaoh hung
Like night and darkened all the land of Nile:
So numberless were those bad angels seen
Hovering on wing under the cope of hell
‘Twixt upper, nether, and surrounding fires;
Till, as a signal givn, the uplifted spear
Of their great Sultan waving to direct (PL, I.338-48)

The biblical story of a tyrant leads to the contemporary realities of the sultans. The "numberless" host of "bad angels" are like the overwhelming army of the Turkish sultans invading Christendom. The images of kingship run through
Paradise Lost and Paradise Regained driving home the central theme of the indictment of the fraudulent kings of history and the glorification of the kingdom of Heaven. Stevie Davies (1991, p. 29) sums this up as follows:

It is through the reiterated image of the king of this world that Milton embodied in Paradise Lost this cyclical pattern of corruption in human life. Once Satan has been established as the archetype of vitiated kingship...the image is extended through a kind of family or chain of subsidiary kings invading all time and all space, from Moloch to Pharaoh to Charles I.

For a Europe embattled with the Ottomans, the image of sultan resonated with menace and dread associations. The aforementioned image of Satan as a sultan at war well conveys these associations; taking into consideration the menace of the Turks as lived experience for Milton’s readers, it was the most chilling image that he could have invoked. Davies (1991, p. 51) notes how the Royalists and the Roundheads in their mutual recriminations called each other the "Turk." For an embattled West these images of the sultan could also be "a test of grace in which readers find themselves. If the West cannot stand against the Sultan, how far can the Christian soul resist Satan?" (Davies, 1991, p. 55) Furthermore, Milton’s universalistic view of history had to include the vituperation of all kings throughout history.

Another important aspect of what we termed "contemporary framework" is the stamp of mercantile and "discovery" accounts collected by Hakluyt (1589) and later Samuel Purchas (1625) in Paradise Lost. As Martin Green (1979, p. 41) maintains, although these accounts were mercantile in character at that time, they were still a form of propaganda providing the "energising myth" of English imperialism; "English propaganda was mercantile in character; and hence it was England that had a Hakluyt and later a Defoe." The following image is another instance of the impingement of mercantilism on the Miltonic text and Milton’s attitude towards it, Satan’s voyage to reach the gate of Heaven is described.
As when far off at sea a fleet descreed
Hangs in the clouds, by equinoctial winds
Close sailing from Bengala, or the isles
Of Ternate and Tidore, whence merchants bring
Their spicy drugs: they on the trading flood
Through the wide Ethiopian to the Cape
Ply stemming nightly toward the pole. So seemed
Far off the flying fiend: (PL, II.636-43)

Is the association of mercantilism with the Satanic regime indicative of Milton’s criticism and disapproval of, in Tony Davies’ words, "the revolutionary energies of expansive colonial capitalism?" (1992, p. 252) Paul Stevens (1997, p. 8) argues that, as with colonialism, Milton only disapproves of mercantile excess.

In the most extensive study of the idea of colonialism in Paradise Lost, Martin Evans discusses the poem’s corollary of orientalist discourse: colonial discourse. He pinpoints the ways in which the poem on a cosmic stage re-enacts many of the salient themes of colonial encounters and conquests. In Evans’ words, "Milton’s version of the genesis myth resonates with the complex thematics of Renaissance colonialism"(1996, pp. 4-5). These texts were, then, concerned, apart from the internal politics, with politics in general including the colonial politics and the East-West encounters manifested in the matter of the Orient. Orientalism and colonialism, in Said’s account, are supposed to be consciously or unconsciously in collusion. However, in its explicit form the colonial discourse in these texts mostly focuses on America and Africa, where colonial activities were well under way. It is implicit with regard to the Orient as references to the colonial stance of England apropos the Orient, specifically India, were not manifest at Milton’s time and were in large part subsumed under a mercantile discourse. Orientalist discourse of the poem, nonetheless, in depicting the East as a locus of untold riches and drawing on and interrogating the discourses of mercantilism, exploration and discovery also negotiated with the ideology of European imperialism. As Evans (1996, p. 11)
points out, the first and crucial phase of English empire-building, which was in the New World, was more or less coincident with Milton’s life-time. Did Milton approve or disapprove of this ideology?

Here we encounter a number of different critical attitudes. Walter Lim (1993, p. 30) is of the opinion that owing to his disillusionment with the English nation Milton interrogates England’s colonial and imperial aspirations but he does so in "poetic narratives controlled by a dominant rhetoric of theological imperialism." David Quint (1993, p. 25) sees *Paradise Lost* as an "indictment of European expansion and colonialism that includes his own countrymen and contemporaries." For Paul Stevens, however, Milton is no anti-colonialist; he does not criticise colonialism per se but only its abuses. His is an ideal colonialism, it is Milton’s "very virtue, his desire for civility and his refusal of any thoroughgoing relativism, that makes it so difficult for him to stand outside the discourse of colonialism" (1997, p. 17). The upshot of all this is that Milton’s attitude towards the ideology of colonialism is split. According to Evans, there are two colonial narratives in *Paradise Lost*, an anti-colonial one, based on the Spanish conquests in the New World, and a pro-colonial one based on the English attempts to settle Virginia and New England. Evans (1996, pp. 141 & 142) traces this bifurcation to the "essentially binary character of English colonial ideology" during most of Milton’s life-time, contending that it fits Stephen Greenblatt’s theoretical model of Renaissance political discourse in which "a 'subversive' critique of diabolic exploration and conquest is 'contained' by a larger history of divine imperialism in which England’s hegemony over its transatlantic possessions is emphatically endorsed." In short, *Paradise Lost* is imprinted with the thematics of colonialism as it is with that of orientalism.

We would call the kind of orientalism with more overt, pronounced imperialist / colonialist assumptions, implications and tendencies "imperialist" orientalism in contrast to a sort of early or "emergent" orientalism which seems to be more "a body of knowledge," an attempt on the part of the West to imagine, map out and finally conceptualise lands and peoples other than itself, the struggle of a culture to come to terms with its oriental others. This of course seems very similar to the distinction
Said (1978, p. 222) makes between "latent" and "manifest" orientalism, the former referring to "an almost unconscious (and certainly untouchable) positivity" and the latter to the "various stated views" about things oriental. The differences in the viewpoints about the Orient in different writers can be "characterized as exclusively manifest differences, differences in form and personal style, rarely in basic content" (Said, 1978, p. 206).

Conclusion

To sum up, in dramatising the conflict, ancient and ongoing, between good and evil, one rich stockpile of ideas, images and figures available to Milton was the matter of the Orient. The oriental images and figures woven into the fabric of the vast epic similes help give aesthetic form to the theological concepts, the ideological economy and the thematics of Paradise Lost. Reiterated, either explicitly or implicitly, throughout this text, specifically as regards the images of kingship, the matter of the East helps unify the symbolics of Milton's epic. The frequent association of the Orient with evil and the satanic regime indicates that, in the final analysis, it cannot be completely wrested from what Said calls "latent Orientalism," which Milton inherits and empowers. But, issues such as aesthetic considerations, disregarded in Edward Said’s account; a cosmic setting and background; drawing on the authority of history and a universalist, Christocentric view of it; the complex negotiation with a multiplicity of other discourses circulating at that time; classicism and an attempt at an encyclopaedic scope; the vilification of all kingship through all time and space and, above all, a profound process of displacement whereby comments on contemporary issues are displaced on to the matter of the Orient all help compound the orientalism of Paradise Lost. Milton's representation of the East, his 'imaginative geography', an instance of the complex interweaving of culture and power, exemplifies the uncanny shaping power of representing the Other on the 'imagination' of a great writer.
References


