



A Socio-cognitive Approach to Developing Oral Fluency and Naturalness in Iranian EFL Learners

Farahman Farrokhi

English Department, Faculty of Persian Literature and Foreign Languages, University of Tabriz, Tabriz, Iran

E-mail: ffarrokhi20@yahoo.co.uk

Asgar Mahmoudi (corresponding author)

PhD Candidate at University of Tabriz, Iran

E-mail: mahmoudi301@gmail.com

Received: 10-08-2013

Accepted: 25-11-2013

Published: 01-03-2014

doi:10.7575/aiac.ijalel.v.3n.2p.1

URL: <http://dx.doi.org/10.7575/aiac.ijalel.v.3n.2p.1>

Abstract

Learning spoken English in situations like Iran that do not support adequate and rich exposure results in English which is slow-paced and reduced regarding all aspects. The purpose of this study was to explore the ways by which these two problems can be accounted for to some extent. The theoretical bases upon which the study was founded were sociocultural and cognitive approaches to language learning. Sociocultural theorists emphasize learners' involvement in social activities and believe that it is enough for learning a language. Cognitive theorists, on the other hand, emphasize the role of memory and the rote learning of instances of language. Adopting an integrative approach, this study employed interaction and rote learning activities in the experimental classes as its independent variables and measured their effects on students' fluency and naturalness, which were the dependent variables of the study. While fluency was defined as the speed of speech and gauged mostly by aggregating the number of syllables produced in the unit of time and the syllable number/phonation time ratio multiplied by one hundred minus the negative values assigned to pauses, naturalness was defined as fluency plus formulaicity. Formulaicity scores were measured by assigning forty points to each formulaic expression produced by participants. Findings from integrated interaction and rote learning activities in the experimental classes were compared with findings from the control class in addition to being compared with each other. The performances of participants in the experimental classes were compared with each other because the rote-learned materials were offered to them in different formats. One of the experimental classes received decontextualized formulas with their meanings in Persian while the other class received contextualized formulas without meanings. Students in the control class were required to reproduce oral texts of short movies in lieu of memorizing formulas. Findings from the study revealed that while interaction alone is enough for developing fluency, it is not enough for developing naturalness. The experimental class receiving decontextualized formulas with meanings outperformed the other two classes in developing naturalness. The ultimate conclusions reached were of two types. First, getting involved in interaction is sufficient for developing fluency, because there was no significant difference among groups with regard to this variable, but it is not enough for developing high levels of naturalness. Second, to reach acceptable levels of naturalness, participants need to memorize formulas whose meanings are provided for them.

Keywords: Idiom/formula, Fluency, Naturalness, Pause

1. Introduction

1.1. General background

In what is called the post-method era, much of the researchers' time and energy has been spent on finding an eclectic method, or what Kumaravadivelu (2003) calls "principled pragmatism" (p. 33). Post-method era is significant largely because of the emergence of Communicative Language Teaching (CLT). CLT, however, is not equally effective in different places largely because of the type and amount of input available to learners. Most students in low-intensity, low-exposure foreign language situations (Randall, 2007) after a few years of schooling in English are able to produce grammatically correct sentences but cannot speak as fast and as idiomatically as natives can. Reduced and laborious language is the hallmark of foreign language learning in many places.

The need for foreign language learners to develop fluency and idiomaticity/formulaicity, in addition to linguistic competence, is beyond dispute today. Skehan (1998) states that, to obtain the automaticity that this involves requires frequent opportunities to link together the components of utterances so that they can be produced without undue effort. Iwashita, Brown, McNamara and O'Hagan (2008), similarly, in examining the nature of speaking proficiency in English, found that vocabulary and fluency have the strongest effect on examiners' impression about participants' proficiency.

1.2. Statement of the problem and purpose of the study

The situation of EFL in Iranian context is completely different from most EFL situations in other parts of the world. In Iranian context, because of cultural and political issues, learners of English as a foreign language have almost no access to native speakers of English for communication neither in real world nor in cyber world and their exposure to live face-to-face English is limited almost only to their non-native instructors. Therefore, the notion of global “target-language community”, as suggested by Harmer (2008, p. 11), has less applicability to this context. The problem with this kind of language learning is that learners are exposed to a kind of English which is unnatural. English spoken by Iranian EFL learners is slow-paced and full of contrived English sentences, the phenomenon that Lewis and Hill (1985) call language-like behavior.

Since, a lot of situations in our lives are recurrent and we follow certain routines in making reference to similar situations, the reliance on word-based invented language instead of chunk-based formulaic and ritualized language frequently used by natives, foregrounds the oddness in the English produced by Iranian EFL students. Flowerdew and Miller (2005), explaining the notion of schema, point out that knowledge is organized and stored in memory according to re-occurring events. They observe that “[o]nce the structure of an event is stored as a schema in memory it aids individuals in negotiating future events, in allowing them to predict what is likely to happen. In a similar way, knowledge of previous texts (spoken or written) also aids in negotiating subsequent texts” (p. 26).

Regarding the problem highlighted above, adopting a mixed method of interaction and memory-based mechanisms with a focus on formulas seems to be what is necessary for Iranian EFL learners to approach the natural and fast flow of speech observed among native speakers of English. To this end, the methodology suggested in this study used insights from sociocultural theory by bringing it down to the level of interaction as its major practical import and incorporated it with rote-learned formulas or exemplars considered important by memory-based information-processing or computational perspectives to language learning. However, formulas were offered to the experimental groups in two different ways: decontextualized with meanings and contextualized without meanings. Therefore, the Interaction-plus-Decontextualized-Formula Group (IDFG) and Interaction-plus Contextualized-Formula Group (ICFG) represented the experimental groups. The final outcomes of these two groups were then compared with the outcome of the Interaction-Plus-Movie Group (IPMG), which constituted the control group. The outcomes of the two experimental groups were also cross-compared to find out which level had best materialized the targeted goals.

1.3. Significance of and justification for the study

Kumaravadivelu (2006) sees it the responsibility of the teacher to help learners reach a desired level of linguistic and pragmatic knowledge/ability. On the other hand, Thornbury (2012) points out that “the ability to *speak* the second language, as opposed to writing or reading it, is typically a priority for most students” (p. 198). These two citations taken together mean that spoken language must be given priority and at the same time taught correctly. It is unfortunate that, neither of the requirements is fully satisfied in Iranian EFL context. And even if the first one is addressed, there are serious doubts as to Iranian English teachers’ pragmatic knowledge/ability. In other words, most instructors are not competent enough to deliver their instructions in natural English.

The use of unnatural and bizarre constructions is just one side of the coin. Another side is the problem of low speed with which Iranian EFL learners speak. The speed of speech in a word-based language system cannot match that of exemplar-based which operates from the procedural memory and uses big extended chunks of language at every attempt. The two problems combined create a sense of inadequacy which is unjustified considering the amount of time and energy that Iranian EFL learners invest on the task of learning English.

The inefficiency of remedial interactive programs, too, is due to the fact that this kind of interaction most of the time occurs between people whose English is reduced and impaired, in the sense explained above, and full of inventions which are far removed from natural and routinized English. The major source of input for Iranian EFL learners is their own interlanguage not the non-simplified authentic language produced by natives, reflecting a point referred to by Kumaravadivelu (2006).

Ohta (2000) states that “clearly the nature of effective assistance in the ZPD varies depending upon a variety of factors, including the expertise of the helper” (p. 76). This pushes us to conclude that the benefits conceived of for interaction in sociocultural theory, are not directly applicable to Iranian EFL context. In situations like Iran, where access to natural communication is denied EFL learners, it seems the methodology of interaction should be seasoned by an exemplar-based methodology focused mainly on the expressions which are normally and routinely used by natives. This is where cognitive approaches and their corollaries, frequency and memory, step in. However, the dilemma is finding the way by which we can bring the two opposing theories together. Ellis (2004) speculates that there are two possibilities to reconcile the computational and socio-cultural perspectives: developing a general theory that incorporates both approaches or accepting the inevitability of theoretical pluralism. Musumeci (2009), too, states that clearly, it would be absurd to reduce our current understanding of second language acquisition to simple dichotomies. In dealing with a complex knowledge system like language, one would expect that certain theories better explain some aspects of the phenomenon than others. In line with Ellis and Musumeci, who tend to accept theoretical pluralism, the purpose of this research is not to pore over the differences in order to develop a general theory of language acquisition. Rather, it tries to investigate the possibility of incorporating the potentials of both theories to find a solution for the kind of problems outlined above.

1.4. Research hypotheses

The null research hypotheses guiding this study are:

- Participants will gain nothing in terms of fluency and naturalness in IPMG, IDFG, and ICFG from pre-tests to post-tests.
- IPMG will be as successful as IDFG and ICFG in developing oral fluency.
- IPMG will be as successful as IDFG and ICFG in developing naturalness.

1.5. Definition of key terms

Idiom/formula

Idiom or formula in this study referred to all kinds of multi-word constructions in English language from phrasal verbs, collocations, clause structures to full sentences. The only criterion for a combination of words to be considered an idiom or a formula was its more than chance occurrence.

Fluency

Fluency referred to the speed of speech and was measured by counting the number of syllables articulated in the unit of time plus the number of syllables/phonation time ratio multiplied by 100 (SPR x 100) minus the values calculated for dysfluency markers (silent and filled pauses and substitutions).

Naturalness

Naturalness was used to refer to the number of idiomatic or formulaic expressions that a student used in the five-minute time limit available to him or her in each recording session. It was measured by counting the number of identifiable formulas (as defined in section 3) and assigning 40 points to each of them. No score was given, however, to repeated formulas or if the number of formulas exceeded one fortieth of the number of syllables produced.

Pause

Pause in this study referred to three things: silence, repetition, and substitution or reformulation. Silence was defined as a period of time longer than three seconds in which the student said nothing. Meaningless gap fillers like *ah*, *er*, etc. were not considered to be words. Repetition meant the exact verbalizing of the previous word or words. And finally, substitution referred to reformulations or situations in which a student changed his or her mind and substituted the already articulated word or words for other alternatives. Substitutions were either complete or partial. All single-word substitutions were naturally complete. But, multi-word substitutions were of both types: partial and complete.

2. Review of the related literature

2.1 Introduction

For some researchers learning is inferential but for others, it is not. Inferential means that we observe the outcomes of learning not learning itself. But, from countless researches it has become evident that many things are influential in learning a second language (Brown, 2007). A categorization of these factors reveals that the major issues are stacked at the two poles of mind and society. There are theorists who give priority to mind at the expense of society (e.g., Segalowitz, 2003) and theories which prioritize society deemphasizing mind (e.g., Ohta2000).

2.2 Sociocultural theory

Sociocultural theory has been the latest influence on SLA. "The most fundamental concept of sociocultural theory is that the human mind is mediated" (Lantolf, 2000, p. 1). Mediation typically takes the form of *assisted performance* within the *zone of proximal development*. Swain and Lapkin (1998) call the opportunities that arise for learning within the ZPD as *occasions for learning*. Within the ZPD, as it is usually conceived of, a more competent interlocutor interacts with the learner to provide a supportive discourse framework. Tharp and Gallimore (1988) call this kind of supportive talk as *instructional conversation* which is, goal-directed and jointly constructed teacher-learner discourse that replicates the reciprocity and contingency of casual conversations. However, communication has a somewhat different meaning in sociocultural theory. For example, Lantolf and Thorne (2006) stress that communication has a place in activity theory if only ontological priority is given to it as a social practice. They reject the idea of communication as sign per se.

vanLier (2004) defines sociocultural theory as a general approach to the human sciences whose goal is to explain the mental functioning and the cultural, institutional, and historical situations relationships. In a similar vein, Ohta (2000) stresses that in sociocultural theory a learner is neither the processor of input, nor the producer of output, but as speaker/hearer involved in developmental processes which are realized in interaction.

"Sociocultural theory is a theory of mind, based on Vygotsky's belief that the properties of mind can be discovered by observing mental, physical, and linguistic activity, because they are *intrinsically* related" (Roebuck, 2000, p. 80). Vygotsky (1978) saw consciousness as a process through which people construct their environments and dynamically organize and realize higher mental functions such as voluntary attention, voluntary memory, intention, planning, and the resulting behavior. Vygotsky (1986) advanced the notion of psychological tools, the most important of which is language, in order to explain human consciousness and introduced the concepts of private and inner speech as mechanisms by which individuals regulate their own behavior.

2.3 Cognitive theory

Cognitive psychology is an interdisciplinary venture that draws upon the insights of psychologists, linguists, computer scientists, neuroscientists, and philosophers to study mind and mental processes (Stillings et al., 1995). Cognitivism is the position that complex mental processes play an important role in shaping human behavior. From a cognitive perspective, learning is an intermental phenomenon inferred from what people say and do. A central theme is the mental processing of information, its construction, acquisition, organization, coding, rehearsal, storage in memory, and

retrieval or nonretrieval from memory. It may be interesting to know that, although cognitive theorists stress the importance of mental processes in learning, they disagree over which processes are more important (Schunk, 2012). Mitchell and Myles (2004) define cognitivists as people who view SLL as one instantiation of learning among many others.

In this paradigm, language learning is seen as the acquisition of a kind of complex procedural skill which is practiced and integrated into fluent performance (McLaughlin, 1987). This requires the automatization of component sub-skills. Without automatization no amount of knowledge will ever translate into the levels of skill required for real life use (DeKeyser, 2001).

In connectionist/emergentist theory of language acquisition, which is just one of many cognitive theories, the emphasis is on usage. Learning does not rely on an innate module, but rather it takes place based on the extraction of regularities from the input. Usage-based theories (also called item-based or exemplar-based theories), according to Dörnyei (2009), constitute a group of related linguistic approaches to the understanding of (primarily first) language acquisition and processing.

In addition to automatization or proceduralizing, frequency, practice, restructuring, and memory are also very important for cognitive theorists. Language for N. Ellis (2009), for example, is estimation from sample. Bybee (2002) also confirms that a wide range of contemporary linguistic approaches recognize that linguistic knowledge is based firmly on language experience, and frequency of use. People progress from a stage of declarative knowledge to a stage of procedural knowledge (vanPatten & Benati, 2010). The role of short-term memory in this transition is vital because, as Carroll (2008) puts it, many cognitive processes require that we hold onto information for a short period of time to be able to commit it to the long-term memory.

2.4 Formula

Definition

According to Walsh (2010), a feature of spoken language is a category of vocabulary often referred to as fixed and semi-fixed expressions, or variously termed chunks, clusters, lexical bundles, idioms, and multi-word units. Formulas and lexical phrases are other terms that are used in many books and papers to refer to these multi-word structures (e.g., Nation & Meara, 2010; Schmitt & Carter, 2004). Recurrent sequences such as *I don't know*, *all of a sudden*, *all over the place*, *don't have a clue* are also variously referred to as n-grams, prefabs, and lexical bundles (Flowerdew, 2008).

These different terms embrace the notion of rote-learned or imitated chunks of unanalyzed language, available for learner use without being derived from generative rules (Myles, Hooper, & Mitchell, 1998). As Pawley and Syder (1983) state, native speakers do *not* exercise the creative potential of a generative grammar to anything like their full extent. Maximally rapid intelligibility is afforded by the use of frequent, pre-existing chunks in the parole (N. Ellis, 2001). Boers and Lindstromberg (2005), too, state that while rule-based language instruction is a well-meant attempt at making language learning more cognitively economical, it has lately been contended because many naturally occurring phrases fall outside the scope of such rules.

Definitions of formulaic sequences center on the notion that they are multi-word units of language that in spontaneous speech production are stored in and retrieved from long-term memory as if they were single lexical units. According to Wood (2006), formulaic sequences are fixed combinations of words that have a range of functions and uses in speech production and communication, and seem to be cognitively stored and retrieved by speakers as if they were single words. They can facilitate fluency in speech by making pauses shorter and less frequent, and allowing longer runs of speech between pauses.

Prevalence

Schmitt and Carter (2004) state that formulas are used in a wide variety of ways. For example, they express concepts (*put someone out to pasture*), state a commonly believed truth or advice (*a stitch in time saves nine*), provide phatic expressions which facilitate social interaction (*Nice weather today*), signpost discourse organization (*on the other hand*), and provide technical phraseology which can transact information in a precise and efficient manner (*Blood pressure is 140 over 60*). Formulaic sequences are a major component of almost all types of discourse both in terms of degree and scope of usage (Axwell, et al., 1998). And finally, as Biber, et al. (1999) explain, due to the pressure of online production, spoken language tends to consist of self-standing phrases and clauses, most of them prefabs, to avoid the syntactic complexity and undue subordination of written language.

Wray (2002) believes that a serviceable core vocabulary will include fixed and semi-fixed, multiword phrases, also known as formulaic language. Current approaches emphasize that authentic sources, especially of native-speaker usage, embody more idiomaticity and because many formulaic expressions have identifiable pragmatic functions, they constitute a core component of the speaker's pragmatic competence (Thornbury, 2012).

Research suggests that at least one-third to one-half of language is composed of formulaic elements (Conklin & Schmitt, 2008; Erman & Warren 2000). Corpus linguistic research, also, demonstrates that natural language makes considerable use of recurrent multiword patterns (N. Ellis, et al., 2008). Biber, et al., (1999) found that 3- and 4-word lexical bundles made up 28 percent of the conversation and 20 percent of the academic prose they studied. Howarth (1998) found that 31-40 percent of academic texts was made up of collocations and idioms. Estimates by Erman and Warren (2000) stood around fifty percent.

Types

Formulas can be divided to open and closed (Grant & Bauer, 2004; Sinclair, 1991). At one end, we have collocations that are in habitual company. At the other end, we have open collocations, which according to Grant and Bauer (2004) are the loosest kind of Multi-Word Utterances (MWUs). Grant and Bauer call the most restricted MWUs core idioms. Liontas (2008), too, defines idiom in a very restricted way. For example, he excludes compound nouns like *man-of-war*, phrasal verbs like *to give in*, prepositional verbs like *to look after*, logical connective prepositional phrases like *for instance*, formulaic expressions like *at first sight*, etc. from the category of idioms.

However, the accepted idea today is to look at the idiom phenomenon or MWUs as an all-embracing continuum of open to closed formulas and in terms of co-occurrence not in terms of compositionality or non-compositionality. Wright (1999), for example, rejects the traditional view of idioms by stating that there is a lot more language which is *idiomatic*. From Aitchison's (2012) perspective too, collocational links cover a wide spectrum and the boundary between idioms and phrases is hazy.

Processing

There must be a reason why formulaic sequences are so widespread. Corrigan et al., (2009) quote Bannaard and Lieven (2009) as arguing that formulaic language occurs because humans show preferences for things they have experienced previously. According to Pawley (2009) prefabricated schemas underpin fluent and idiomatic speech. Finally, Conklin and Schmitt (2008) provide a sociofunctional explanation for the pervasiveness of formulaic sequences, but they believe that there is a psycholinguistic explanation as well. The psycholinguistic explanation, according to Dörnyei (2009) is that, formulaic sequences are stored in the memory as single units and therefore their retrieval is cognitively relatively undemanding.

2.5 Fluency

According to Leaver, Ehrman, and Shekhtman (2005), the amount of information you have does not determine your level of fluency; what is important is what you do with your knowledge. The ultimate goal of many second-language (L2) learners is to be fluent in the target language and express their thoughts easily in any given situation (de Jong & Perfetti, 2011). But, people usually accumulate a lot of up-in-the-head knowledge about language and then find that they cannot actually use this language to communicate when they want to. A reason for this can be a lack of experience which according to Scrivener (2011) may make learners feel nervous about saying things. Fluent language users have had tens of thousands of hours on task. They have processed many millions of utterances involving tens of thousands types presented as innumerable tokens (N. Ellis. 2001).

The first clear definition for fluency was indeed suggested by Fillmore (1979) based on temporal phenomena. Fillmore also pointed out that fluency is closely related to a couple of other things including speaker's lexical knowledge, knowledge of fixed linguistic forms, and the knowledge of formulaic expressions. Brumfit (1984) felt that fluency is to be regarded as natural language use which gives speech natural and normal qualities like native-like use of pausing, rhythm, intonation, stress, rate of speaking, etc. Today, fluency is regarded a phenomenon which is largely dependent on practice and proceduralization (Anderson, 1995; Ellis, 2001; MacLaughlin, 1991). Logan's (1989) instance theory is an alternative to these skill-building approaches. This exemplar-based approach suggests that increased automaticity is the function of a growing repertoire of memorized bits of domain-specific knowledge (labeled instances). Once the repertoire of memorized instances has been built up, the learner does not have to rely on following rules but can retrieve a relevant stored instance as a single step.

Measuring fluency

Wood (2006) gives a brief chronological description of the efforts in examining speech fluency. Wood's description shows that it was Goldman Eisler who first looked at the temporal variables of speech or articulation rate (measured as syllables uttered per minute or second), length of fluent runs, and pause phenomena like length and frequency in 1967. Foster, Tonkyn, and Wigglesworth (2000) introduced the measure of focusing on units or chunks of spoken language within lengthy turns. They introduced macro- and micro-planning processes. The former may cover quite long (e.g. multi-sentence) stretches of speech, and the latter shorter units similar to clause or sentence.

Iwashita et al.(2008) in an attempt to find the degree of fit between language proficiency and scores assigned by raters to spoken samples measured fluency with regard to filled pauses, unfilled pauses, repair, total pausing time (as a percentage of total speaking time), speech rate, and mean length of run. Three aspects of pausing are usually measured in fluency research: duration, frequency, and syntactic location. Studies have shown that pause times in L2 speech are generally longer than in L1 speech more so because of longer planning time or monitoring mistakes (Scovel, 2001). Gries (2008) used frequency lists to quantify and/or compare the attainment of language proficiency. According to this researcher, the most basic and most frequent statistic employed in this context is the type/token ratio.

2.6 Naturalness or native-likeness

For some researchers native-likeness is a vague and elusive term (e.g., Davies, 2006; van Lier, 2004). Many others believe that achieving a native-like mastery is impossible (Bley-Vroman, 1990; Johnson & Newport, 1989). However, there are researchers who do not see the problem of second language acquisition far from insuperable. For example, Dörnyei (2009) claims that there is accumulated evidence which leaves little doubt that some late starters can master an L2 to an extent that they would be regarded by most people as native speakers of the particular language. There are researchers, too, who do not reject the possibility of becoming native-like but believe that it is a formidable task which is dependent on the fulfillment of some conditions. For example, (Ellis, 2009) believes that extensive exposure is

necessary for native-like selection. Many of the forms required for idiomatic use are relatively low frequency, and the learner thus needs a large needs-relevant authentic input sample to encounter. More usage is still required to allow the tunings underpinning native-like use of collocation – something which even advanced learners have particular difficulty with.

3. Methods

3.1 Participants

The participants of this study were all freshmen students (three classes) majoring in Teaching English as a Foreign Language (TEFL) at Mohaghegh Ardabili University's Namin branch, in the northwest town of Namin, Iran. The students had taken up a four-credit conversation course whose aim was to develop the students' communication skills. The course was the students first ever conversation course in a university. Each class consisted of more than 25 students, three to four of them male students. The age range was quite small (18-22) excluding a couple of older female students in each class. The students were from different ethnic backgrounds and different parts of the country.

3.2 Sampling

The sampling procedure for selecting the participants involved two stages. In the first stage the entire populations of students enrolled for the courses were included in the study but in the second stage students were screened based on information from their pretests for the sake of variance homogeneity. The screening process was carried out following a study by Farrokhi and Mahmoudi (2012). To form the most comparable groups a number of conditions and criteria were defined for students' eligibility and exclusion from the study. The factors controlled for in this study were age, years of schooling in English, syllable number per unit of time (SYL), phonation time (PT), syllable number/phonation time ratio (SPR), silent pauses, filled pauses, and substitutions or reformulations. The ranges defined for these factors are given in Table 3.1. Not to violate the assumptions of statistical tests, enough care was exercised to limit participants' attributes in a way that allowed the researchers to have enough and, at the same time, almost equal numbers of participants in each group.

Table 3.1. Ranges Defined for Obtaining Variance Homogeneity

FEATURE	RANGE
Age	18-22 years old
Years of schooling in English	2-4 years
Phonation time in pretest	250-300 seconds
Number of syllables in pretest	400 and above
SPR	1.40 and above
Number of silent pauses	0-5
Number of filled pauses	0-30
Number of substitutions	0-25

3.3 Setting

The classes were held twice a week in Namin, a small town in northwest Iran with each session lasting for an hour and a half. The university in this town is a division of a main university called Mohaghegh Ardabili University located in the provincial city of Ardabil. All class meetings were held in a language laboratory. When the researcher started teaching, the laboratory's new facilities, including computers and monitors and a central control system for the instructor had just been installed, but, unfortunately, the software needed for the system to work was not installed yet. Therefore, the equipment was only used for recording the students' voices by the Movie-maker Programs accompanying Windows XPs installed on the computers. Regular class meetings were held on the empty side of the lab so that students could work in groups or whole class format.

3.4 Instrumentation

The researcher decided to use monologues in the pretests and posttests for a variety of problems associated with subjective tests like interviews. To objectively choose the students who were eligible for the study and to have pretest and posttest measures whose reliability could be verified with high precision, a battery of three tests each including 21 topics was designed. The topics were the same in all three versions of the test but their arrangements were different. The students were free to speak about whichever topic they felt comfortable with. There was no limit also set on the number of topics to speak about, but the students were required to fill the given five-minute time by speaking and produce texts longer than a single sentence for each topic. The reliability of this test was calculated as soon as the post-test was given to the control group. The Pearson's r turned out to be above .80 for all essential pairs of measures correlated (including syllable number, PT, SPR, and fluency scores) indicating considerably high temporal consistency.

Cronbach alpha is one of the most commonly used indicators of internal consistency. Cronbach alpha, however, is sensitive to item number and finds quite low correlation values with small sample sizes. For this reason, a correlation coefficient was calculated for five main scales (SYL, PT, SPR, fluency score, and naturalness score) by incorporating participants' scores in all three groups in the pretests and a Cronbach alpha value was calculated. Table 3.2, represents the obtained value.

Table 3.2 Cronbach Alpha Reliability Coefficient

Cronbach's Alpha	N of Items
.772	15

The validity of the instrument was also established, even though the validity of a scale is determined to some extent by its reliability if its content has consistency with the goal(s) for which it is designed. Determining validity is an empirical process and involves collecting evidence for tests' use and efficiency. In the case of oral tests, recording students' output for subsequent analyses in different situations is a well-documented practice (e.g., Candlin & Mercer, 2001; Hellermann, 2008). Recording of the type used in this study, like other studies using recording as their instrument of data collection, provided continuous access to students' oral productions so that carrying out in-depth analyses on what might have been otherwise impossible, with regard to the fleeting nature of spoken language, became possible. Moreover, the advantage of this procedure was that it avoided the polluting factors that might have crept in if the recordings had not been in the form of monologues.

3.5 Procedure

The data for the study were collected in the two consecutive semesters in 2012 during spring and fall and lasted for four months each semester including five sessions of staging period. The students had been enrolled for the conversation courses that were offered by the university. Data collection procedures were the same in all three control and experimental classes although the activities that students engaged in were somewhat different. In the IPMG, the students were given 60 short movies each of about 2 to 4 minutes long to watch (three for every session) and were tested randomly each session for their ability to reproduce the texts in the movies. The movies were downloaded from different weblogs or picked up from Britannica and Encarta encyclopedias. Watching movies was an adjunct to interaction, which was the main activity, and was given as a placebo to the control group to keep time on interaction roughly constant for all groups.

Students in the IDFG had received a pamphlet containing 1500 formulaic expressions with their meanings in Persian printed in front of them. These students were required to memorize 100 formulaic expressions every session up to session ten, excluding six sessions prior to the beginning of the experiment, and 50 formulaic expressions from session ten to session twenty. The reason for reducing the number of expressions to be memorized after the tenth session was to decrease the work load and provide students with enough time for reviewing. Like students in the control group, these students were also tested randomly every session to see if they had mastered the introduced formulaic expressions and their meanings.

The third class or ICFG was handled in a similar way. That is, the pamphlet given to the students included the same formulaic expressions, but it differed in that the meanings of expressions were not supplied. Instead, the formulas were embedded within sentences and, in a sense, were contextualized. The students had to explore the meanings of the expressions by themselves.

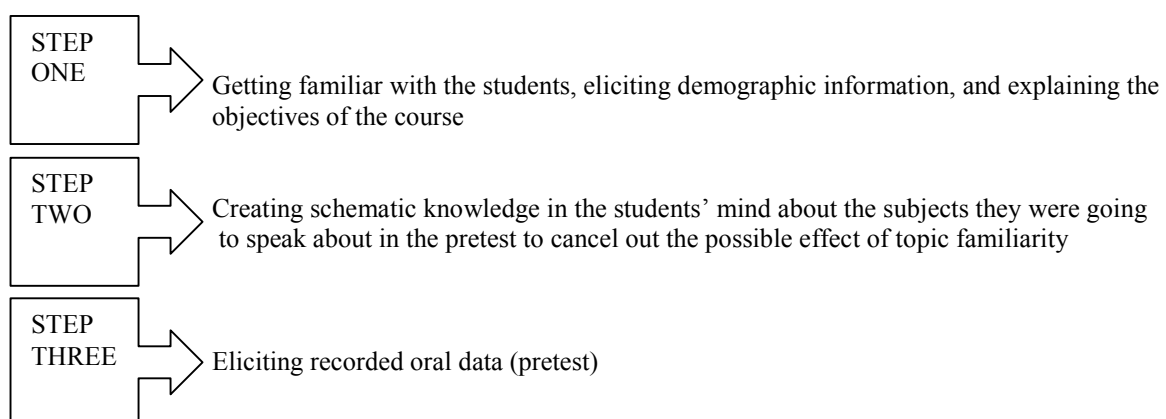
The methods employed with experimental groups were theoretically motivated by Laufer and Hulstijn (2001). In the context of vocabulary learning, most studies, according to Hulstijn (2001), refer to the *depth of processing*. Following Laufer and Hulstijn (2001), the current researcher tried to keep the cognitive loads of the tasks given to experimental classes roughly equivalent. That is, both tasks were +need, +search, and +evaluation. Since the need for performing both tasks was externally imposed, they were +need. In the case of search, one group was required to look for the structures in which the idioms or formulas could possibly be used. The meanings for idioms were provided for this group. The other group, on the other hand, was supposed to look for the meanings of the formulas provided for them in the context. Therefore, both tasks' search loads were moderate. The same held true for evaluation loads, i.e., they were moderate, because every idiom or formula either in isolation or within the sentence had only one meaning.

Selecting formulas for teaching is a very sensitive issue and cannot be done only by considering their frequency profiles. There are a few reasons for this. First, some formulas are so frequent that there is no need for teaching them like *How are you? I think, For example, and In fact*. Second, some formulas, like proverbs, are very infrequent but known to almost all people. Third, cultural issues should be taken into consideration. Some idiomatic expressions are frequent among natives but may be out of place in situations like Iran as is the case with *high as a kite* which means drunk or a *foxy woman* which means a sexy woman. Fourth and finally, some idioms are used by select groups of people like addicts. For example, *He needs another fix* which means he needs another dose of drug.

To make these differentiations and select appropriate formulas, the researcher decided to appeal to the intuitions of two experienced conversation instructors. The 1500-formula collection was the result of five sessions in which the researcher and the instructors put their thoughts together and selected intuitively the most frequent but at the same time appropriate formulas. The number 1500 was decided on for feasibility considerations.

3.6 Steps taken in IPMG

Teaching in the control class comprised of a series of steps taken by the researcher. Ignoring the details of the activities, the overall flow of the teaching is represented in Figure 3.1.



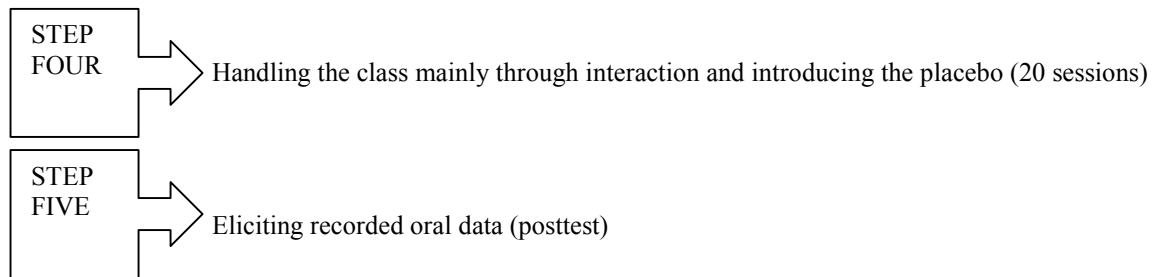


Figure 3.1. Summary of the Steps Taken in IPMG

3.7 Steps taken in IDFG and ICFG

Steps taken for data collection and the number of sessions held in IDFG and ICFG were exactly the same as the steps taken and the number of sessions held in the IPMG. The treatments given to these classes, however, differed in that students in IDFG had to rote learn decontextualized formulaic expressions which were provided for them alongside their meanings and students in ICFG were to rote learn the same formulaic expressions by extracting their meanings from the sentence-level contexts in which they were embedded. The interaction or conversation phases of the classes which followed each phase of formula review were designed to emphasize and consolidate the use of these idiomatic expressions in actual speech. Students were given credits for the idiomatic or formulaic expressions they attempted or incorporated in their conversations but the scoring procedure was holistic and was employed only after each phase of conversation was finished not to interrupt the natural flow of students' speech. It is important to note that, the researcher tried his best to employ the same set of formulas in his interactions with students in IPMG over and over again by commenting on the students' points of view, but the focused and conscious practice of formulas, characteristic of experimental groups, was absent in this class.

3.8 Design and variables of the study

This study was conducted based on a Mixed Between-Within Factorial design with two independent and two dependent variables. It was mixed because students' gains in fluency and naturalness were measured and compared within the groups themselves from the pretests to posttests as well as between the groups. It was factorial because there were two dependent and two independent variables. The schematic representation of the design of the study is given in Figure 3.2, below.

Independent Variables	Dependent variables	
	Fluency	Naturalness
Rote learning of formulas		
Interaction		

Figure 3.2 Design of the Study

The two independent variables of this study were interaction (in the form of group and whole-class discussions) and rote learning of formulas (decontextualized formulas with meanings and contextualized formulas without meanings). The effect of interaction was decided to be measured by counting the number of syllables produced by participants in the pretests and comparing the obtained values with the number of syllables produced in the posttests. The effect of rote learning of formulas was determined by calculating and comparing the formulaicity scores (the number of formulas multiplied by 40) for each student in pretests and posttests. The dependent variables of the study were fluency and naturalness. Fluency meant the speed of speech and was measured by a variety of scales, to be discussed below, whereas naturalness was used to refer to fluency plus formulaicity.

Regarding the measurement of dependent variables, fluency literature reveals a variety of methods by which this construct is measured. Inspired by previous works and following Farrokhi and Mahmoudi (2011), a variety of scales including substitutions, silent and filled pauses, SYL, PT, and SPR were used to measure the participants' fluency in both pretests and posttests. This multi-faceted measurement strategy was adopted because some measures are more sensitive than others in measuring fluency. For example, counting syllables is more sensitive than counting words, because words are constituted of one or more syllables. Or, SPR represents the speed of speech better than the number of syllables per minute because it takes account of silent pauses and eliminates their effect on the speed of speech.

The Farrokhi-Mahmoudi model was originally designed to find best or least matching pairs and convergent and divergent groups. The model was used with some modification in this study, however, to measure students' fluency. The procedure involved attributing every student an overall number for his or her fluency which was calculated in a series of stages as follows:

1. Calculating the number of syllables for each student and arranging them from the highest to the lowest,
2. Calculating the students' SPR, multiplying them by one hundred and finally arranging the numbers from the biggest to the smallest,
3. Counting the number of silent pauses in each student's output and assigning five negative points to each of them and then arranging the numbers from the biggest to the smallest,

4. Tallying the number of repetitions or filled pauses for each student's production, attributing three negative points to each of them, and then arranging the overall negative scores in a descending order, and finally
5. Counting the number of substitutions for each student, assigning one negative point to each of them, adding them up and arranging them from the biggest to the smallest.

The stages can be summarized in the following formula:

Number of syllables + (SPR x 100) + (number of silent pauses x -5) + (number of filled pauses x -3) – number of substitutions

It is clear that the bigger the two first measures the more fluent the person will be considered to be but the last three measures will have an aggravating effect on the person's fluency score because these measures amount to negative values. The calculations of fluency scores for two students from the study are given in Table 3.3, for more illustration.

Table 3.3 Calculations of Fluency Scores

	Student A	Student B
SYL	654	661
SPR	$2.32 \times 100 = 232$	220
Silent pause score	$4 \times -5 = -20$	0
Filled pause score	$4 \times -3 = -12$	-36
Substitution score	$1 \times -1 = -11$	-4
Fluency score	843	841

The reason why SPR is multiplied by one hundred is that, this is the most indicative factor of participants' speed of speech. The number of syllables shows a participant's overall production of syllables only. It does not tell us about the effect of possible pauses. SPR, in contrast, reveals the actual speed by which the person has been articulating the syllables while speaking, because pausing time effect is eliminated by deducting it from the denominator of the fraction which represents the phonation time (PT).

The augmentation of SPR, of course, should happen in a way that gives us a number that cannot be exceeded by the negative points. In Table 3.1, it can be seen that the limits set on and the ranges defined for silent pauses, filled pauses, and substitutions do not allow speakers' negative scores to overtake -140. Since the minimum SPR determined for participants in order to be included in the study was 1.4, even if a student produced the maximum number of silent and filled pauses and substitutions, allowed by the design of the study, the resulting negative number would not exceed the lowest SPR multiplied by 100, which gives 140. Once this number is determined, using bigger numbers is pointless because they will not change the order of outcomes. For example, if we multiply the SPRs in Table 3.3, by 120 the obtained numbers for students A and B will be 278 and 264, respectively. The fluency scores, in this case will be 889 for student A and 885 for student B. As it is evident, the difference increases by two points only and does not change the order of fluency. However, if SPR were multiplied by a number smaller than 100, there would be the danger of a change in fluency order because of some students' excessive pauses. In the above example, multiplying student A's SPR by 50 reduces the acquired number to 116 while the same number for student B turns out to be 110. Adding up the numbers in this case gives the fluency score of 727 for student A and 731 for student B. As it can be seen, the fluency order is reversed, but it is clear enough from students' SPRs that student A has been speaking faster.

The values for other measures of fluency are also decided based on some logical reasoning, although arbitrarily. Silent pauses are the most serious dysfluency markers. Filled pauses fall in the second place and are more serious in their own right than substitutions which fall in the third place. Silent pauses show a speaker's lack of strategic competence in addition to other disqualifications. While those who fill their pauses by repeating themselves suggest a higher level of fluency and thwart the implication that they are not capable of expressing ideas. Regarding substitutions, it is clear that, most of the time, their function is clarifying.

Now that we have settled the problem of calculating fluency, it is time to suggest a method for calculating naturalness. As it was said above, naturalness embraces the concepts of fluency or speed of speech and formulaicity. We have already suggested a way for calculating fluency. This solves half of the problem. For the other half, it was decided to use two principles: *the introduced formulas principle* and *the one-to-forty principle*. The first principle says, formulas are the expressions introduced by the researcher to the experimental groups. This principle will be applied to pretests and posttests alike to be able to trace the appearance of formulas in the students' posttest productions and make uniform judgments. The reason for applying this principle was that formulas are so widespread in any text that identifying all of them is almost impossible. According to the second principle, each formulaic expression has the value of forty syllables in spoken speech. This criterion was chosen based on the ratio of syllable number to identifiable formulaic expressions in the outputs of five more advanced students in the pilot study. Since from the perspective of this study, speed of speech and formulaicity both are important for the perception of naturalness, it was decided that roughly equal importance be given to both fluency and formulaicity. Therefore, if a student's fluency score, for example, were 800 and he or she produced 10 idiomatic or formulaic expressions, his or her overall naturalness score would be 1200. On the other hand, this student would not be given credit for more than 20 formulaic expressions, because as the scarcity of idiomatic expressions in a student's output may result in unnaturalness, excessive use of these expressions, too, can result in anomaly. No credit was also decided to be given to the repetitions of formulaic expressions, because a person may repeat a formulaic expression like *for example* in his or her speech lots of times.

4. Data analyses

4.1 Preliminary analyses

A description of groups' and participants' selection criteria in the pretests was given in the previous section. The descriptive statistics of all three groups with regard to all attributes measured were also calculated in pretests. The randomness issue was dealt with in section 2. However, a non-normal distribution would endanger the validity of a study. The following three tables represent the results of three tests of normality for the groups each including tests of normality for three important attributes measured in the pretests. These three attributes, henceforth scales, were essential for determining participants' fluency and naturalness levels as will be discussed later.

Table 4.1 Tests of Normality for IPMG

	Kolmogorov-Smirnov ^a		
	Statistic	df	Sig.
pretest syllables number	.137	16	.200*
pretest syllable number phonation time ratio	.183	16	.158
pretest fluency score	.113	16	.200*

Table 4.2 Tests of Normality for IDFG

	Kolmogorov-Smirnov ^a		
	Statistic	df	Sig.
pretest syllable number	.167	15	.200*
pretest syllable number phonation time ratio	.142	15	.200*
pretest fluency score	.164	15	.200*

Table 4.3 Tests of Normality for ICFG

	Kolmogorov-Smirnov ^a		
	Statistic	df	Sig.
pretest syllable number	.176	16	.200*
pretest syllable number phonation time ratio	.182	16	.160
pretest fluency score	.230	16	.023

All three sets of normality tests evidently show values which are not significant. Therefore, we can conclude that the distribution of scores for measured scales in our samples were normal. Significant differences among groups' distributions of scores regarding the measured scales (lack of homogeneity of variance), too, can potentially distort the results of the study and make it difficult to attribute the findings to interventions. Leven's Tests of homogeneity showed that this was not an issue to concern us. To complete the preliminary analyses section and to be able to test the null hypotheses of the study, the descriptive statistics of the participants' performances in the posttests were also calculated.

4.2. Testing the first null hypothesis

To test the first null hypothesis, we had to run Paired-Samples T-test for each group regarding the dependent variable in question since all our data were continuous and the distributions were normal. Tables 4.4, 4.5, and 4.6, indicate that students' gains in all three groups in relation to both fluency and naturalness had been statistically significant from time 1 to time 2.

Table 4.4 Gains in Fluency and Naturalness from Pretest to Posttest in IPMG

	Paired Differences						t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error	95% Confidence Interval					
				Lower	Upper				
Pair 1 pre fluency post fluency	-82.18750	86.85637	21.71409	-128.46999	-35.90501	-3.785	15	.002	
Pair 2 pre naturalness post naturalness	-169.68750	131.28479	32.82120	-239.64423	-99.73077	-5.170	15	.000	

The calculated effect sizes were $D=.49$ and $D=.64$ for these two variables which in the case of fluency was very close to moderate, but in the case of naturalness it was moderate.

Table 4.5 Gains in Fluency and Naturalness from Pretest to Posttest in IDFG

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error	95% Confidence Interval				
					Lower	Upper			
Pair 1	pre fluency score post fluency score	-80.53333	102.22375	26.39406	-137.14296	-23.92371	-3.051	14	.009
Pair 2	pre naturalness score post naturalness score	-280.53333	220.22971	56.86307	-402.49248	-158.57419	-4.933	14	.000

The calculated effect sizes for the obtained significance levels in IDFG were $D=.40$ and $D=.63$ for these two variables which in the case of fluency was somewhat smaller than the value calculated for IPMG, but in the case of naturalness was almost the same.

Table 4.6 Gains in Fluency and Naturalness from Pretest to Posttest in ICFG

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error	95% Confidence Interval				
					Lower	Upper			
Pair 1	pre fluency post fluency	-91.25000	130.54322	32.63581	-160.81157	-21.68843	-2.796	15	.014
Pair 2	pre naturalness post naturalness	-215.00000	221.99219	55.49805	-333.29129	-96.70871	-3.874	15	.001

The calculated effect sizes in ICFG for the obtained significant levels were $D=.34$ and $D=.48$ for these two variables. Both of the effect sizes calculated for ICFG, though not very small, are smaller than the effect sizes calculated for IDFG and ICFG.

Although findings from the T-tests reject the first null hypothesis, comparing the three pairs of effect sizes may catch the readers by surprise finding out that the effect sizes calculated for the control group's fluency and naturalness are bigger than the experimental groups' effect sizes. However, in closer inspection this is not the case thoroughbred. Running Paired-Samples T-tests for SPR and formulaicity scores in each group in pretests and posttests, which were the major factors in determining fluency and naturalness respectively, and calculating their effect sizes revealed this.

Table 4.7 T-test Results for SPR and Formulaicity Scores in Pretests and Posttests (IPMG)

		Mean	Std. Deviation	Mean	t	df	sig
Pair 1	pretest SPR posttest SPR	-.16375	.21515	.05379	-3.044	15	.008
Pair 2	pretest formulaicity posttest formulaicity	-87.50000	85.44004	21.36001	-4.096	15	.001

Table 4.8 T-test Results for SPR and Formulaicity Scores in Pretests and Posttests (IDFG)

		Mean	Std. Deviation	Mean	t	df	sig
Pair 1	pretest SPR posttest SPR	-.22133	.38221	.09869	-2.243	14	.042
Pair 2	pretest formulaicity posttest formulaicity	-205.33333	152.59033	39.39865	-5.212	14	.000

Table 4.9 T-test Results for SPR and Formulaicity Scores in Pretests and Posttests (ICFG)

		Mean	Std. Deviation	Mean	t	df	sig
Pair 1	pretest syllable phonation ratio posttest syllable phonation ratio	-.14875	.30824	.07706	-1.930	15	.073
Pair 2	pretest formulaicity posttest formulaicity	-117.50000	152.11837	38.02959	-3.090	15	.007

Using the same formula for calculating the effect sizes, we had $D=.38$, $D=.26$, and $D=.20$ for fluency in IPMD, IDFG, and ICFG, respectively. The naturalness effect sizes in the same order were $D=.53$, $D=.66$, and $D=.39$. From these values we can understand that while the effect sizes calculated for fluency in all groups were modest, of course, with the biggest value belonging to IPMG, the effect size calculated for naturalness in IDFG was bigger than the values calculated for the same measure in IMFG ($.66 > .53$) and ICFG ($.66 > .39$). These conflicting results can be resolved only by attributing the effect sizes for fluency to groups' sameness of treatment for fluency, i.e., interaction and different treatments for naturalness, i.e., rote learning.

4.3 Testing the second and third null hypotheses

Paired-Samples T-tests used in testing the first null hypothesis revealed substantive gains in terms of fluency and naturalness for all three groups from pretests to posttests. However, to figure out if the gains in these two dependent variables had been uniform across groups (to test null hypotheses 2 and 3) we needed to perform One-way between groups ANOVA for each dependent variable's posttest results and compare all three groups' means. One-way between groups ANOVA is used when there is one independent grouping variable, i.e., with more than two levels, and one dependent variable. The independent variable can also be a continuous variable that has been recoded to give two or more equal groups, as in the case of our study. A non-significant value for this test will tell us that the gains across groups have not been significantly different. A significant p-value, however, will point to a difference somewhere in the data. In case a difference is identified, a post-hoc test will be needed to spot the location of the difference.

Like other parametric tests, One-way between-groups ANOVA has some assumptions. These assumptions are:

1. The dependent variable is measured at the interval or ratio level;
2. The scores are obtained using random samples from the population;
3. The observations are independent. That is, the behavior of one participant or group does not affect the behavior of other participants or groups or is not influenced by them;
4. The populations from which the samples are taken are normally distributed.
5. The samples are obtained from populations of equal variances.

The first assumption was met because all scales in this study for both dependent and independent variables were continuous. The Randomness issue was discussed in section 2 and resolved by indicating that the participants had no chance for self-selection. The third assumption was also met because the testing procedure guaranteed the independence of observations by individually testing the participants both in pretests and posttests. There was also a ten-weeks-long time interval between pretests and posttest which was enough to consider that there was no test effect. The normality of the population, which is the subject of fourth assumption, was also verified and proved by running the Kolmogorov-Smirnov tests and obtaining non-significant p-values (see Tables 4.1, 4.2, and 4.3.). And finally, the homogeneity of variances of populations from which the samples were obtained were checked by running the Leven's test as part of analysis of variances (ANOVA) and the values obtained were non-significant with $p=.647$ for fluency and $p=.485$ for naturalness pointing to the non-violation of this assumption.

Tables 4.10, and 4.11, represent the results of this kind of analysis of students' gains in fluency and naturalness. The important point about these tests, however, is that while the p-value in Table 4.10 is not significant, it is significant in table 4.11. The p-value obtained from the first test, therefore, cannot reject our second null hypothesis. In other words, while our participants in all three groups had made substantial gains in terms of fluency from pretests to posttests (determined by T-tests), the mean of their gains were not significantly different from each other in posttests. In contrast, from the second test, we understand that the participants' gains in naturalness in the control and experimental groups were not uniform and there was a significant difference. This makes running a post-hoc test inevitable to find out where the difference exactly lied. Table 4.12, which follows the ANOVA tables, shows the results of Scheffe post-hoc test.

Table 4.10 One-way between Groups ANOVA Comparing Posttests Resultsfor Fluency

postfluency	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	96415.617	2	48207.809	1.991	.149
Within Groups	1065312.808	44	24211.655		
Total	1161728.426	46			

Table 4.11 One-way between Groups ANOVA Comparing Posttests Resultsfor Naturalness

postnaturalness	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	470018.628	2	235009.314	5.397	.008
Within Groups	1915918.308	44	43543.598		
Total	2385936.936	46			

Table 4.12 Multiple Comparisons of Means for Posttest Naturalness Scores
Dependent Variable: posttest naturalness

	(I) group	(J) group	Mean Difference		
			(I-J)	Std. Error	Sig.
Scheffe	1.00 IPMG	2.00 IDFG	-239.45417*	74.99588	.010
		3.00 ICFG	-66.12500	73.77635	.672
	2.00 IDFG	1.00 IPMG	239.45417*	74.99588	.010
		3.00 ICFG	173.32917	74.99588	.080
	3.00 ICFG	1.00 IPMG	66.12500	73.77635	.672
		2.00 IDFG	-173.32917	74.99588	.080

Looking down the Mean Difference column in Table 4.12, we can see that some values are marked by asterisks. This means that the two compared groups had been significantly different at $p < .05$ level. A quick glance at the table tells us that while IPMG/ICFG and IDFG/ICFG differences had not been significant, IPMG/IDFG difference had been significant in terms of naturalness. This finding brings us to the point of rejecting the third null hypothesis.

That the difference between IDFG and IPMG's posttest results in terms of naturalness had been significant, however, does not tell us whether this difference had been large enough. To understand the strength of difference, we have to calculate the effect size of the difference using multivariate analysis and checking the estimates of effect size box from the options menu. The analysis gives the value of $D = .968$ which is a very strong effect size, close to perfect.

5. Discussion and conclusions

Of the three research hypotheses in this study one of them was ultimately rejected while two were accepted. The accepted hypotheses propose two things:

1. All three groups in the study (one control and two experimental) made significant gains with respect to fluency from pretests to posttests.
2. The gains of IDFG were greater than the gains of ICFG with respect to naturalness.

The rejected hypothesis, however, makes it clear that the gains of IDFG and ICFG were not significantly different from each other in relation to fluency.

These results confirm that interaction and social contexts are essential for developing spoken language ability and in a way support the sociocultural dimension of the study. Learning speaking and developing fluency constitute a major part of learning language. However, learning to speak a language is not equal to learning natural spoken language. The second hypothesis of the research addresses largely the naturalness dimension of spoken language and supports a cognitive approach to language learning. This finding suggests that a practice or rote learning phase of formulas should be incorporated into any program of teaching spoken language. Nevertheless, it should be kept in mind that in our operationalizing of naturalness, fluency was considered as part of the construct naturalness. Thus, from the perspective of the current researchers the social aspect of language learning is primary while memorizing of formulas plays an ancillary role in improving the quality of speech.

On the other hand, we learned from the research that the availability of formulas is not the only needed condition. What is important is the accessibility of formulas to students. The students in our decontextualized formulas with meanings group (IDFG) found formulas easier to learn than our students in the other experimental group to whom the meanings of formulas were not provided. This, however, does not mean that students will not develop a degree of naturalness if they are not fed formulas with their meanings attached to them. Some students are quick to pick up on and develop acceptable degrees of naturalness, but for the majority of students this becomes a very long and never-ending journey. This brings our discussion round to our original claim that social and cognitive activities should both be encouraged in language classrooms, the former for learning language and developing fluency and the later for naturalizing the learned language.

Acknowledgement

It should be noted that the study reported here is part of a more general doctoral project to be presented to the English department at the University of Tabriz..

References

- Aitchison, J. (2012). *Words in the mind*. Sussex: Blackwell Publishers Ltd.
- Anderson, J. R. (1995). *Learning and memory: An integrated approach*. Oxford: John Wiley & Sons Inc.
- Axwell, K. et al. (1998). *Cambridge international dictionary of idioms*. New York: Cambridge University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Longman.
- Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In S. M. Gass & J. Shachter (Eds.), *Linguistic perspectives on second language acquisition*. (pp. 41–69). Cambridge: Cambridge University Press.
- Boers, F., & Lindstromberg, S. (2005). Finding ways to make phrase-learning feasible: The mnemonic effect of alliteration. *System*, 33(2), 225–238.
- Brown, H. D. (2007). *Principles of language learning and teaching* (5th ed.). New York: Pearson Education Inc.
- Brumfit, C. (1984). *Communicative methodology in language teaching: The roles of fluency and accuracy*. New York: Cambridge University Press.

- Burns, A. (2012). Text-based Teaching. In A. Burns & J. C. Richards (Eds.), *Pedagogy and practice in second language teaching*. (pp. 140–148). New York: Cambridge University Press.
- Bybee, J. (2002). Phonological evidence for exemplar storage of multiword sequences. *SSLA*, 24(2), pp. 215–221.
- Candlin, C. N., & Mercer, N. (2001). *English language teaching in its social context*. London: Routledge.
- Carroll, D. W. (2008). *Psychology of language*. Belmont: Thomson Corporation.
- Conklin, K., & Schmidt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, 29(1), 72–89.
- Corrigan, R., Moravcsik, E. A., Ouali, H., & Wheatley, K. M. (2009). Introduction: Approaches to the study of formulae. In R. Corrigan, E. A. Moravcsik, H. Ouali & K. M. Wheatley (Eds.), *Formulaic Language* (1st Volume). (pp. XI–XXIV). Amsterdam: John Benjamins Publishing Company.
- Davies, A. (2006). The native speaker in applied linguistics. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics*. (pp. 431–450). Oxford: Blackwell Publishing Ltd.
- De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning*, 61(2), 333–368.
- DeKeyser, R. M. (2001). Automaticity and automatization. In Robinson, P. (Ed.), *Cognition and second language instruction*. (pp. 125–151). Cambridge: Cambridge University Press.
- Dörnyei, Z. (2009). *The psychology of second language acquisition*. New York: Oxford University Press.
- Ellis, N. C. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction*. (pp. 3–32). Cambridge: Cambridge University Press.
- Ellis, N. C. (2009). Optimizing the input: Frequency and sampling in usage-based and form-focused learning. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching*. (pp. 139–159). Oxford: Blackwell Publishing Ltd.
- Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, 42(3), 375–396.
- Ellis, R. (2004). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition*. (pp. 63–104). Victoria: Blackwell Publishing Ltd.
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Interdisciplinary Journal for the Study of Discourse*, 20(1), 29–62.
- Farrokhi, F., & Mahmoudi, A. (2011). An innovative way of finding best or least matching pairs or groups. *English Language and Literature Studies*, 1(2), pp.129–141.
- Farrokhi, F., & Mahmoudi, A. (2012). Rethinking convenience sampling: Defining quality criteria. *Theory and Practice in Language Studies*, 2(4), pp.784–793.
- Fillmore, C. J. (1979). On fluency. In C. J. Fillmore, D. Kempler & W. S. Y. Wang (Eds.), *Individual differences in language ability and language behavior*. (pp. 85–101). New York: Academic Press.
- Flowerdew, J., & Miller, L. (2005). *Second language listening: Theory and practice*. New York: Cambridge University Press.
- Flowerdew, J. (2008). Corpora in language teaching. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching*. (pp. 327–351). Oxford: Blackwell Publishing Ltd.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(3), 354–375.
- Grant, L., & Bauer, L. (2004). Criteria for re-defining idioms: Are we barking up the wrong tree? *Applied Linguistics*, 25(1), 38–61.
- Gries, S. T. (2008). Corpus-based methods in analyses of second language acquisition data. In P. Robinson & N. C. Ellis (Eds.), *A handbook of cognitive linguistics and second language acquisition*. (pp. 406–432). New York: Routledge.
- Harmer, J. (2008). *How to teach English*. Essex: Pearson Education Ltd.
- Hellermann, J. (2008). *Social actions for classroom language learning: New perspectives on language & education*. Clevedon: Multilingual Matters Ltd.
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19(1), pp. 24–44.
- Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics*, 29(1), 24–49.
- Johnson, J., & Newport, E. (1991). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60–99.
- Kumaravadivelu, B. (2003). *Beyond methods: Macrostrategies for language teaching*. New Haven: Yale University Press.
- Kumaravadivelu, B. (2006). *Understanding language teaching*. New Jersey: Lawrence Erlbaum Associates.
- Lantolf, J. P. (2000). Introducing sociocultural theory. In J. P. Lantolf (Ed.), *Sociocultural theory and second language acquisition*. (pp. 1–27). New York: Oxford University Press.
- Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural Theory and the genesis of second language development*. New York: Oxford University Press.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–26.
- Leaver, B. L., Ehrman, M., & Shekhtman, B. (2005). *Achieving success in second language acquisition*. New York: Cambridge University Press.
- Lewis, M., & Hill, J. (1985). *Practical techniques for language teaching*. London: Language Teaching Publications.
- Liontas, J. I. (2008). Toward a critical pedagogy of idiomaticity. *Indian Journal of Applied Linguistics*, 34(1), pp. 11–27.

- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological review*, 95(4), 492–527.
- McLaughlin, B. (1987). *Theories of second language learning*. New York: Routledge, Chapman and Hall, Inc.
- Mitchell, R., & Myles, F. (2004). *Second language learning theories*. New York: Oxford University Press.
- Musumeci, D. (2009). History of language teaching. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching*. (pp. 42–63). Oxford: Blackwell Publishing Ltd.
- Myles, f. Hooper, J., & Mitchell, R. (1998). Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning*, 48(3), 323–363.
- Nation, P., & Meara, P. (2010). Vocabulary. In N. Schmidt (Ed.), *An introduction to applied linguistics*. (2nd ed.). (pp. 34–53). New York: Oxford University press.
- Ohta, A. S. (2000). Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning*. (pp. 51–79). Oxford: Oxford University Press.
- Pawley, A., & Syder, F. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. Richards & R. Schmidt (Eds.), *Language and communication*. (192–225). London: Longman.
- Pawley, A., (2009). Grammarians' languages versus humanists' languages. In R. Corrigan, E. A. Moravcsik, H. Ouali & K. M. Wheatley (Eds.), *Formulaic Language (1st Volume)*. (pp. 3– 27). Amsterdam: John Benjamins Publishing Company.
- Randall, M. (2007). *Memory, psychology, and second language learning*. Amsterdam: John Benjamins Publishing Co.
- Roebuck, R. (2000). Subjects speak out: How learners position themselves in a psycholinguistic task. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning*. (pp. 79–97). Oxford: Oxford University Press.
- Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmidt (Ed.), *Formulaic sequences: Acquisition, process and use*. (pp. 1–23). Amsterdam: John Benjamins Publishing Co.
- Schunk, D. H. (2012). *Learning theories: An educational perspective*. Boston: Pearson.
- Scovel, T. (2001). Psycholinguistics. In R. Carter & D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages*. (pp. 87–93). Cambridge: Cambridge University Press.
- Scrivener, J. (2011). *Learning teaching: The essential guide to English language teaching*. New York: MacMillan.
- Segalowitz, N. (2003). Automaticity and second languages. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition*. (pp. 382–409). Victoria: Blackwell Publishing Ltd.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Stillings, N. A., Weisler, S. E., Chase, C. H., Feinstein, M. H., Garfield, J. L., & Rissland, E. L. (1995). *Cognitive science: An introduction*. Cambridge: MIT Press.
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The modern Language Journal*, 82(3), pp. 320–338.
- Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning and schooling in a social context*. New York: Cambridge University Press.
- Thornbury, S. (2012). Speaking Instruction. In A. Burns & J. C. Richards (Eds.), *Pedagogy and practice in second language teaching*. (pp. 198–207). New York: Cambridge University Press.
- vanLier, L. (2004). *The ecology and semiotics of language learning*. Massachusetts: Kluwer Academic Publisher.
- vanPatten, B., & Benati, A. G. (2010). *Key Terms in Second Language Acquisition*. New York: Continuum International Publishing Group.
- Vygotsky, L. S. (1978). *Mind in Society: The development of higher psychological processes*. Harvard: By the President and Fellows of Harvard University.
- Vygotsky, L. S. (1986). *Thought and language*. (Newly revised & edited by A. Kozulin). Cambridge: The MIT Press.
- Walsh, S. (2010). What features of spoken and written corpora can be exploited in creating language teaching materials and syllabuses? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics*. (pp. 333–345). New York: Routledge.
- Wood, D. (2006). Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. *The Canadian Modern Language Review*, 63(1), 13–33.
- Wray, A. (2002). *Formulaic language and lexicon*. Cambridge: Cambridge University Press.
- Wright, J. (1999). *Idioms organizer*. London: Thomson Learning, Inc.