Exploring L2 Speech-Production Management from a Cognitive Perspective:
A Focus on Fluency

Azizullah Mirzaei a
Assistant Professor, Shahrekord University, Shahrekord, Iran

Najmeh Heidari b
M.A. in TEFL, Shahrekord University, Shahrekord, Iran

Received 4 April 2012; revised 13 August 2012; accepted 22 August 2012

Abstract
In L2 instruction and assessment, the application of research findings that adopt an integrative, psycholinguistic approach to explore the information-processing and speech-management facets of fluency seems necessary. This paper reports on a study that drew on Levelt’s (1989, 1999) speaking-specific model to probe the problem-solving mechanisms (PSMs) of fluent and nonfluent L2 speakers through a speaking-oriented questionnaire and a series of output-related retrospective interviews. The fluent and nonfluent L2 speakers were identified using a newly-developed analytic fluency scale and the task-related speech samples of 200 participants. The results revealed that the fluent L2 speakers employed cognitive, linguistic, and interactional PSMs more frequently and with greater facility than the nonfluent participants particularly to compensate for deficits in their conceptual repertoire, bridge communication gaps, and negotiate the intended meaning with their interlocutors. Specifically, they efficiently reshaped the preverbal plan to avoid failure, adeptly employed a variety of fillers and hesitation devices to

a Email address: mirzaei-a@lit.sku.ac.ir, fazizullah@yahoo.com
Corresponding address: English Department, Faculty of Letters & Humanities, Shahrekord University, 2nd kilometer of Saman Road, Shahrekord (Chaharmahal Bakhteyari Province), Iran. Zip code: 8818634141, Phone number: 0098-9125588577

b Email address: heidari_najmeh67@yahoo.com
maintain the communication flow, and attentively monitored the conversation. However, the nonfluent L2 speakers entirely abandoned or completely changed their original speech plan after running into deficiencies in their own outputs that made their speech utterly disjointed. Still at times, they struggled to self-correct their speech but failed due to deficient linguistic and interactional competence, which adversely led to more disfluencies. The findings suggest that further research into the cognitive, linguistic, and interactional processes underlying (non)fluent speakers’ use of PSMs can be useful in explaining speech disfluencies or learners’ differential L2 fluency.

Keywords: Psycholinguistic approach; L2 fluency; Levelt’s model; Problem-solving mechanisms

Introduction
Describing the underlying cognitive processes and mechanisms that characterize second or foreign language (L2) learning has been the focus of many cognitively-oriented psychologists and psycholinguists (e.g., Dell, 1986; Donald, 1991; Garrett, 1990; Kormos, 2006; Levelt, 1989, 1999; McLaughlin, 1987). One of the main concerns of these cognitive or psycholinguistic researchers has been examining the information-processing and speech-management mechanisms that L2 learners employ while speaking and communicating in either their first language (L1) or L2 (e.g., Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Dussias, 2001; Jared & Kroll, 2001). The available psycholinguistic frameworks of speech production largely focus on language processing in actual speech which is naturally constrained by the human cognitive and psychological processes and resources. These frameworks (including Levelt’s model) attempt to account for and portray the conceptualization, formulation, and articulation processes of speech production.

Levelt (1989, 1999) considered these processing components from the standpoint of automaticity in L1 speech production and claimed that whereas the conceptualizer component is constrained by the attentional resources and requires controlled or conscious processing, the formulator and articulator components operate unconsciously and automatically. However, the related research on L2 speech production has found that the formulating and articulating phases of L2 speech production are executed with conscious awareness, and the proceduralization of linguistic knowledge occurs in the formulator (e.g., de Bot, 1992; Sajavaara, 1987; Towell, Hawkins, & Bazergui, 1996). That is why most L2
learners suffer from a fluency gap and spend a greater deal of time to manage their different processing deficiencies. Obviously then, L2 speakers often consciously or unconsciously employ different problem-solving mechanisms (PSMs) or speech-management strategies to compensate for their possible cognitive, linguistic, and interactional shortcomings (Færch & Kasper, 1983; Tarone, 1977). The body of literature on communication strategies has already shown that fluent and proficient L2 speakers have more competence in using a variety of effective PSMs or strategies to maintain the conversation flow (Dörnyei, 1995; Dörnyei & Scott, 1997). From a psycholinguistic perspective, there are four kinds of PSMs associated with the main problem types of resource deficits, processing time pressure, deficiencies in one’s own output, and deficiencies in the interlocutor’s performance (Dörnyei & Kormos, 1998; Dörnyei & Scott, 1997; Poulisse, 1993).

Researching differentially fluent L2 speakers’ speech-production or speech-management processes and mechanisms from a cognitive perspective can better portray the psycholinguistic resources or processes that are quintessentially involved in oral language processing (Kroll & Sunderman, 2005). For instance, the slower rate of L2 speech processing and production can be attributed to the fact that grammatical and phonological encoding phases demand conscious attention, and L2 speech production is not fully automatic (Raupach, 1987; Towell et al., 1996; Wiese, 1984). This study was an attempt to probe into the psycholinguistic processes or constraints underpinning fluent and nonfluent speakers’ use of PSMs in L2 oral communication. To this end, Levelt’s (1989, 1999) cognitive speaking-specific model was practically drawn upon to explore the processes or operations underlying the four main kinds of PSMs associated with the deficiency sources of resource deficits, processing time pressure, deficiencies in one’s own output, and deficiencies in the interlocutor’s performance.

**Literature Review**

Speaking fluently in either L1 or L2 is a complex cognitive ability that requires the conversion of declarative knowledge into procedural and automatic knowledge. Cognitive processes of L2 production have been extensively attended to through developing psycholinguistic models such as Levelt’s (1989, 1993, 1995, 1999) speech production model, and Anderson’s (1982, 1983, 1993) skill acquisition theory. Among them, Levelt’s (1989, 1999) modular model has been the most commonly cited framework in L2 research to explain the processes underlying L2 speech production. Levelt (1989) believed that speech production is accomplished
by means of a series of complex and interconnected processes such as conceptualizing, formulating, and articulating.

The conceptualizing component generates the communicative intentions and encodes them into a coherent conceptual plan. Levelt (1989) believed that this preverbal message is produced by means of macroplanning and microplanning. Macroplanning specifies the communicative or speech act intentions, and then the microplanning stage decides on adopting an appropriate perspective to convey the message. The preverbal message functions as the input of the formulator. The formulating component consists of lexical entries and retrieves information from the speaker’s mental lexicon. The lexical entries include lemmas that define the meaning and syntax of lexical items and lexemes that contain information about the phonological and morphological features. The formulator transforms the preverbal message into an internal speech which is then transferred into articulator to execute the phonetic plan. At the end, the monitoring phase of speech production perceives the entire process to identify any mistake that may occur in the subprocesses of conceptualization, formulation, and articulation (Levelt, 1999).

Levelt’s (1989, 1999) speaking-specific model has been broadly used to survey the processes involved in L2 speech production (e.g., Bygate, 2001; de Bot, 1992; Doughty, 2001; Izumi, 2003; Towell et al., 1996). This model is drawn upon in this study in an attempt to probe the cognitive processes underlying L2 speakers’ speech production management and their use of PSMs. In fact, L2 speakers need to utilize a variety of PSMs or communication strategies in order to compensate for their underdeveloped mental lexicon as well as their conscious and serial encodings of the grammatical and phonological phases of speech production (Kormos, 2006; Skehan, 2009). In order to avoid taxonomies that are only based on surface phenomena and performance-related aspects of speech production, the PSMs should be equally based on the underlying cognitive and psycholinguistic processes (Bongaerts, Kellerman, & Bentlage, 1987; Kellerman, 1991). Dörnyei and Kormos (1998) drew upon Levelt’s (1989, 1993, 1995) psycholinguistic model to classify the four main types of PSMs related to the main deficiency sources of resource deficits, processing time pressure, perceived deficiencies in one’s own output, and perceived deficiencies in the interlocutor’s performance.

Upon encountering difficulty in planning and encoding the preverbal message, L2 speakers may employ several resource-deficits PSMs such as lexical,
grammatical, and phonological PSMs. First, lexical PSMs help L2 speakers to retrieve the proper lemma that conforms to the concepts determined in the preverbal plan. Second, grammatical PSMs handle L2 speakers’ imperfect grammatical knowledge and the word-ordering rules. Third, phonological PSMs manage difficulties that L2 speakers may encounter in the phonological encoding and articulating of the surface structure (Dörnyei & Kormos, 1998). In addition, processing-time-pressure mechanisms such as fillers and hesitation devices are required for fluent L2 speech production because speech processing in L2 is serial and demands more attention in the grammatical and phonological encoding phases (Kormos, 2006). If the monitor detects a problem in the language output, the L2 speaker may resort to a number of strategies such as self-initiation, self-correction, or self-repair. The final category of PSMs is related to perceived deficiencies in the interlocutor’s performance that lead to meaning-negotiation mechanisms.

The related literature on the L2 learners’ use of strategies or PSMs has reiterated that, in order to become successful, learners need to be (inter)actively and attentively engaged in the learning process and employ a set of compensatory and creative learning strategies and interactive techniques (e.g., Griffith, 2008; Rubin, 1975; Wong & Nunan, 2011). PSMs are process-oriented devices that make learners more autonomous and communicatively competent and endow L2 learners with great facility to compensate their inadequate L2 linguistic and communicative competence. In a related vein, L2 speaking fluency involves the ability to apply one’s linguistic competence fluently and naturally, and the underlying cognitive processing of speech production must operate efficiently and automatically.

Fluency has been predominantly defined as either a synonym of oral proficiency or as an element of speaking proficiency referring to the temporal aspects of speech production (Lennon, 1990, 2000). In order to become a more productive and efficient concept for L2 research, fluency should become independent from general language proficiency. The temporal aspects of speech production are empirically more practical and, thus, more applicable for studying the managerial processes of speech production (Chambers, 1997). In this sense, fluency is considered as a performance phenomenon and is differentiated from other elements of oral proficiency such as idiomaticness, appropriateness, lexical range, and syntactic complexity that are related to linguistic knowledge (Lennon, 1990). Schmidt (1992) also considered fluency as a temporal phenomenon and treated it as an automated procedural skill.
Studies of L2 fluency have operationalized fluency by and large in terms of temporal aspects of speech production such as speech rate, repairs, amount and frequency of hesitation, location of pauses, and length of runs of fluent speech between pauses that are linked to the psycholinguistic facets of performance and production (e.g., Derwing, Rossiter, Munro, & Thomson, 2004; Kormos & Dénes, 2004; Lennon, 1990, 2000; Wood, 2010). Fulcher (1996) suggested that temporal variables in L2 speech are associated with the underlying psycholinguistic mechanisms and developed a fluency rating scale. Fulcher analyzed the kinds of hesitation phenomena that could best predict fluency scores in language proficiency tests, summarized rater interpretations of them, and advanced a new scale of fluency.

In sum, the related literature on fluency reinforces the overall impression that fluency is a fairly intricate notion, covering a number of cognitive, linguistic, and interactional facets (Freed, 1995). Thus, an in-depth understanding of this multifaceted construct necessitates a complementary cognitive approach so that a coherent perspective of the underlying mechanisms of L2 fluency and also the processes of proceduralization or automatization can be presented (Segalowitz, 2010). The knowledge of fluent and nonfluent L2 speakers’ use of the PSMs and the information-processing underlying the use of these PSMs has important theoretical and practical implications for L2 fluency development and assessment. The present study thus aimed at scrutinizing the cognitive processes involved in the fluent and nonfluent L2 speakers’ use of PSMs by drawing upon Levelt’s (1989, 1999) psycholinguistic speaking-specific model.

The Study
There seems to be a paucity of research on the cognitive bases and speech-management facets of L2 fluency as an information-processing and problem-solving operation, and studies of L2 fluency have largely focused on the performance-based aspects of this multidimensional construct. Therefore, this study addressed the following research question to account for the processes involved in the problem-management of the fluent and nonfluent L2 speakers.

- What types of problem-solving mechanisms do fluent and nonfluent Iranian L2 speakers employ to circumvent or mitigate the deterrent effects of their deficiency sources?
Method

Participants
The participants in the main phase of this research included 60 fluent (32 males and 28 females) and 60 nonfluent (29 males and 31 females) L2 speakers who were selected from a pool of 200 L2 learners based on the assessment of their audio-recorded speech samples. They were from both graduate and undergraduate levels (aged 20-30) majoring in English Translation, Literature, and TEFL at several Iranian universities. All of the participants were native speakers of Persian and had not been to English-speaking countries. Before they entered university, they had studied English as part of their general academic curriculum for seven years during junior and senior high schools and had little opportunity to use English language for communicative purposes outside the classroom context. However, the undergraduate students were attending general and technical EFL courses at university approximately four days a week and were receiving their education mainly in English as the medium of instruction. For example, they had conversations, oral presentations, and audio-visual translation courses and, thus, had already benefited from more opportunities to use English for authentic purposes. As to the graduate TEFL students, English was the all-purpose medium of instruction used for in-class discussions, presentations, and their theses.

Instrumentation and Data Collection Procedure
Firstly, to select the fluent and nonfluent L2 speakers, an analytic fluency rating scale was developed by integrating and adapting both Fulcher’s (1996) holistic fluency and the International English Language Testing System’s (IELTS) (2008) speaking scales. Analytic scales are believed to be more beneficial than holistic scales because the score given to each criterion yields diagnostic facts about various aspects of learner performance (Carr, 2000). Moreover, analytic scales give raters more opportunity to concentrate on fewer facets of language in giving a score and, thus, are more reliable (Hamp-Lyons, 1991). The new analytic fluency rating scale was composed of hesitation, repetition and restructuring, circumlocution, coherence (Fulcher, 1996), grammatical range and accuracy, and pronunciation (IELTS, 2008) descriptors. The logic behind selecting these descriptors was that they are connected to the psycholinguistic mechanisms involved in speech production (Lennon, 1990; Möhle, 1984; Towell et al., 1996). In addition, studies of L2 fluency have included these descriptors as basic criteria for assessing L2 fluency (e.g., de Jong & Perfetti, 2011; Fulcher, 1996, 2003; IELTS, 2008; Kormos & Dénes, 2004). The new analytic fluency scale also
consisted of six bands that ranged from zero to five. In other words, an interpretive coding system was developed to describe the effects the aforementioned descriptors on language use and clarify why the occurrence of the phenomena interrupts fluency only in certain circumstances (Fulcher, 1996). For example, raters may not consider some pauses as communication impasses but as thinking time for recalling the content of the subsequent expression. In order to warrant the validity and practicality of the new analytic fluency scale, the choice of the descriptors and the scale-development process received constant expert consultation and an intensive revision process.

The participants were required to perform two tasks in both monologic and dialogic conditions after they were allowed to do short pre-task planning for each. Their speech samples were collected by means of a high-quality digital audio-recorder in a quiet room. In the first task, the participants were asked to narrate a six-frame set of cartoons sequenced logically for about two-three minutes. The picture description task is practically beneficial in a research setting because it permits researchers to confine the speech of the participants, meanwhile, does not infuse certain words or structures into their speech (Segalowitz, 2010). The second task was a dialogue in which two participants discussed the benefits and detriments of the media such as the Internet, TV, or satellite. In order to ensure inter-rater reliability, five standardization meetings were held prior to rating the speech samples. Then, two raters, an assistant professor and a student of TEFL, familiar with measures of fluency and the adapted analytic fluency rating scale as well as Levelt’s (1989, 1999) model rated the audio-recorded speech samples. Moreover, Kappa measure of agreement was run to make sure about the raters’ consistency in rating the recordings. The Kappa value was .742 \( (p < 0.05) \), indicating a good estimate of inter-rater consistency.

In the next stage, 60 fluent and 60 nonfluent L2 speakers were selected using the developed analytic fluency rating scale and their audio-recorded speech samples. The participants’ use of PSMs was then inspected through administering a PSM-oriented questionnaire (see Appendix) and conducting a series of output-related retrospective interviews. The questionnaire was developed based on a fairly profound review of the relevant literature on L2 fluency, the descriptions proposed for the PSMs (e.g., Dörnyei & Kormos, 1998; Dörnyei & Scott, 1997), and Levelt’s (1989, 1999) speaking-specific model as well as benefiting from constant consultation with experts (including W. J. M. Levelt, personal communication,
August 31, 2012). It is important to note that the questionnaire was original and creative in the L2 fluency assessment field. The questionnaire was assessed on a four-point Likert scale ranging from never (1) to always (4). The content validity of the instrument was taken care of by developing and using a detailed item specification as the blueprint, experts’ judgments, and pilot testing to ensure that the instrument was precisely and carefully devised to represent a practical sample of L2 speakers’ use of PSMs in oral communication.

The construct validity of the test was analyzed through running factor analysis (Principal Component Analysis). The results yielded satisfactory item loadings on the four main components of the test and, thus, supported the use of the instrument for exploring L2 speakers’ use of PSMs. The reliability of the instrument was estimated using the Cronbach’s alpha that was found to be 0.89, indicating good internal consistency. At the end, retrospective interviews were conducted focusing on the output-related speech samples of 20 participants randomly selected from both groups. The rationale for carrying out retrospective interviews was to detect the fluent and nonfluent L2 speakers’ intentions of using certain PSMs and also to identify the particular subtypes of the PSMs they used to manage their speech production and hereby overcome or mitigate the stumbling effects of the deficiencies they encountered in their task-based L2 oral communication.

**Results**

This study was designed to survey the PSMs the Iranian fluent and nonfluent L2 speakers utilized to surmount communication impasses and convey their intended message. Descriptive statistics were first computed for the PSMs of the participants. Then, a one-way between-groups multivariate analysis of variance (MANOVA) was performed to examine if the differences between the fluent and nonfluent L2 speakers’ use of PSMs was statistically significant. Furthermore, retrospective interviews were conducted randomly sampling participants in both groups. The results of descriptive statistics for the fluent and nonfluent L2 speakers’ use of PSMs are displayed in Table 1. Before moving on, it is important to note that the minimum and maximum scores for each subtype of the PSMs were 1 and 4, respectively.
Table 1 shows that the fluent L2 speakers’ mean scores for the use of PSMs ranged from 2.22 to 3.44 and the nonfluent L2 speakers’ mean scores from 2.23 to 3.32. The results of the descriptive statistics related to the resource-deficits PSMs revealed that the fluent L2 speakers (M = 3.44, SD = .26) employed lexical PSMs more frequently than the nonfluent ones (M = 2.59, SD = .22). However, the nonfluent L2 speakers (M = 3.27, SD = .41) used grammatical PSMs more often than the fluent participants (M = 2.60, SD = .45), and the nonfluent participants (M = 3.31, SD = .52) also employed phonological PSMs more than the fluent participants (M = 2.22, SD = .64). As to the processing time pressure PSMs, it was found that the fluent (M = 2.46, SD = .60) and nonfluent (M = 2.52, SD = .72) L2 speakers employed nonlexicalized pauses almost equally, and also the fluent (M = 2.24, SD = .62) and nonfluent (M = 2.23, SD = .66) participants used repetitions to the same degree. In contrast, the fluent participants (M = 3.29, SD = .49) made use of lexicalized pauses more regularly than the nonfluent ones (M = 2.24, SD = .64). Concerning PSMs associated with perceived deficiencies in own output, the findings indicated that the nonfluent L2 speakers (M = 3.32, SD = .39) utilized self-correction mechanisms more often than the fluent ones (M = 2.65, SD = .40), but
the fluent participants ($M = 2.65, SD = .40$) used check-questions slightly more than the nonfluent participants ($M = 2.23, SD = .64$). The results related to perceived deficiencies in the interlocutor’s performance showed that the fluent participants ($M = 3.22, SD = .33$) employed meaning-negotiation strategies more regularly than the nonfluent participants ($M = 2.69, SD = .52$).

To see if there were statistically significant differences between the fluent and nonfluent L2 speakers with regard to the PSMs they employed to avoid or mitigate the stumbling effects of their deficiency sources, a MANOVA was run. The continuous dependent variables involved in the analysis were PSMs related to resource deficits (i.e., lexical, grammatical, and phonological PSMs), processing time pressure (i.e., nonlexicalized pauses, lexicalized pauses, and repetitions), perceived deficiencies in one’s own output (i.e., self-corrections and check-questions), and perceived deficiencies in the interlocutor’s performance (i.e., meaning-negotiation PSMs). The categorical independent grouping variable was fluency (fluent and nonfluent L2 speakers). Preliminary assumptions of normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicollinearity were checked, and no serious violation was noted.

The results of multivariate tests of significance revealed that the difference between the fluent and nonfluent L2 speakers on the combined dependent variables reached a statistical significance, $F(9, 110) = 82.71, p < 0.0005$; Wilks’ Lambda = .129; partial eta squared = .87. This finding suggests that the Iranian fluent and nonfluent L2 speakers utilized specific PSMs to evade the effect of different cognitive, linguistic, and interactional deficiencies they encountered in L2 communication and that their differential use of PSMs might be an indication of their different deficiency sources or their differential L2 fluency. In order to examine which PSMs they employed differently and, thus, which deficiencies they differently experienced, the results of the tests of between-subject effects are exhibited in Table 2. In this study, to reduce the chance of type 1 error, a Bonferroni adjusted alpha level of .005 was used.
Table 2 indicates that the difference in the fluent and nonfluent L2 speakers’ use of lexical PSMs, $F(1, 118) = 356.22, p < 0.0005$, partial eta squared = .75, was statistically significant. An inspection of the mean scores revealed that the fluent L2 speakers benefited more from lexical PSMs (e.g., all-purpose-words, literal translation, and approximation) in oral communication. Further illustrations of the use of these PSMs will be provided below as part of the in-depth analysis of the speech-samples and the related retrospections. Similarly, the results pertinent to the processing-time-pressure PSMs indicated that the fluent L2 speakers employed lexicalized pauses, $F(1, 118) = 100.58, p < 0.0005$, partial eta squared = .46, more significantly than the nonfluent participants. Importantly, the difference between the fluent and nonfluent L2 speakers’ use of meaning-negotiation strategies, $F(1, 118) = 42.50, p < 0.0005$, partial eta squared = .26, was equally statistically significant with the fluent participants showing greater facility at employing more meaning-negotiation strategies to overcome the out-of-the-blue L2 communication impasses. The fluent L2 speakers typically reported the use of asking for repetition, clarification, and confirmation as well as interpretive summary and guessing subcategories of the meaning-negotiation strategies in the survey instrument. In sum, these results showed that lexical PSMs, lexicalized pauses, and meaning-negotiation strategies were preferred or pragmatically used as more effective PSMs.
by the fluent L2 speakers to circumvent or mitigate the bottle-neck effects of the speech-production deficiencies in communication, whereas for the nonfluent speakers the story was totally different.

Further inspection of the results showed that the nonfluent speakers more frequently sought refuge to grammatical PSMs, $F(1,118) = 69.81, p < 0.0005$, partial eta squared = .37, and phonological PSMs, $F(1,118) = 104.95, p < 0.0005$, partial eta squared = .47, when they felt restrained by their L2-resource-deficits in oral communication. In clear terms, the nonfluent participants more significantly agreed that they typically change the meaning and syntax of a lexical entry through transfer, overgeneralization, or simplified-grammar when encountering deficiencies in L2 grammatical knowledge. In practice, for instance, they may produce a series of incomplete or wrong forms and structures (e.g., it’s kind of) while trying to retrieve and articulate the optimal form of a lexical item. In addition, they demonstrated marked use of self-corrections (instead of trying more social tactics such as meaning-negotiation) to grapple with the perceived deficiencies in their own outputs, $F(1,118) = 84.08, p < 0.0005$, partial eta squared = .41. The nonfluent L2 speakers, specifically, reported more frequent use of PSMs such as error repair, rephrasing repair, and appropriacy repair. In brief, it will be shown through the complementary output-based retrospections below that speech-production problems occurred to the nonfluent speakers’ largely due to their feeling incapable of configuring a viable syntactic structure onto which they could map out their conceptual meanings (if any) or feeling insecure about their phonological competency. This, in turn, engaged them more heavily in resorting to L1-oriented grammatical or phonological PSMs to obviate their L2 inadequacies or in constantly faltering to correct themselves after perceiving deficiencies in their own outputs.

As noted, retrospective interviews were also conducted as an attempt to more deeply examine the (non)fluent participants’ information-processing and speech-management facets of (dis)fluent speech production and, thus, inspire more confidence in the questionnaire-based findings of the study. The qualitative analysis of the retrospective interviews indicated that the fluent participants specifically took advantage of the lexical PSMs such as approximation, all-purpose-words, literal translation, circumlocution, and direct appeal for help. Below are several examples of the fluent L2 speakers’ use of a few of these PSMs. (Pauses less than three seconds are shown with […].)
Episode 1:
Student A: *You should take utmost care to sit on your ... bench appropriately.*
Retrospective comment: I told the word that first came to my mind and used bench instead of chair.
Episode 2:
Student B: *Most programs in satellite are designed to ... what-do-you-call-it certain ideologies in people’s minds.*
Retrospective comment: I forgot the verb inculcate and used what-do-you-call-it to continue the conversation.
Episode 3:
Student C: *I became so addicted to on-line games that sometimes fall asleep at my desk and forget to turn off my ... studying light.*
Retrospective comment: I didn’t notice that I used a studying light instead of study lamp and said the first equivalent came to my mind.
Episode 4:
Student D: *The future generation will be more ... or will be less informative or educated than the other societies in the world.*
Retrospective comment: I couldn’t remember the word illiterate and gave a definition for it.

The above examples clearly illustrate the situations in which the fluent L2 speakers’ needs to devise or use lexical PSMs arose while speaking. In this subcategory of the resource-deficits PSMs, the fluent L2 speakers maneuvered their linguistic knowledge to retrieve the appropriate lemma that fitted their intended meanings. In the first Episode, a fluent L2 speaker used an alternative lexical item (i.e., bench) that shared the semantic features with his planned lexical item (i.e., chair). In other words, he used approximation strategy to convey his message. In Episode 2, when a fluent participant could not remember his intended lexical item, he used a general or an all-purpose word to keep the communication channel open. Episode 3 shows that a fluent L2 speaker translated a lexical item unconsciously resulting in an uninterrupted speech. In the fourth Episode, a fluent participant circumlocuted or described the properties of his intended lexical item. Generally, these examples revealed that the fluent participants were adept at maneuvering their available linguistic knowledge to make their speech more automatic and natural. Furthermore, an inspection of the mean scores together with the retrospective interviews demonstrated that the fluent L2 speakers utilized the fillers subcategory of the processing time pressure PSMs as well as several
subcategories of perceived deficiencies in the interlocutor’s performance such as asking for clarification, asking for confirmation, interpretive summary, and guessing. The following examples illustrate the fluent participants’ use of these effective PSMs.

**Episode 5:**
Student E: Firstly, I believe that TV has made a very very ... you know big gap between the members of a family.
Retrospective comment: I was thinking about the next expressions.

**Episode 6:**
Student F: In my opinion, advertisements have negative effects on children.
Student G: What do you mean?

**Episode 7:**
Student H: I think TV and computers have harmful effects on our body. Especially, those children who spend lots of time watching TV or playing computer games will become fat and lazy.
Student I: So, you are inclined to highlight the negative sides of these media, but I disagree; everyone should learn to use the media appropriately.

The above examples present the fluent L2 speakers’ use of fillers and meaning-negotiation strategies. The fluent L2 speakers specifically showed great facility at using a variety of processing time pressure PSMs such as fillers to sustain the conversation. For example, in Episode 5, a fluent participant used a filling word (i.e., *you know*) to temporize and remember the next expression. Episodes 6 and 7 display the fluent L2 speakers’ use of meaning-negotiation strategies. In Episode 6, a fluent participant asked for clarification and, in Episode 7, another fluent participant gave an interpretive summary of the interlocutor’s message. The overall results point to the fact that the fluent Iranian L2 speakers tended to benefit from a variety of effective PSMs to convey their intended message and overcome communication impasses.

On the contrary, a scrutiny of the mean scores and the qualitative data revealed that the nonfluent L2 speakers employed grammatical and phonological PSMs related to resource-oriented deficits. In addition, they employed a series of self-correction PSMs such as error repair, rephrasing repair, different repair, and appropriacy repair. The participants’ uses of these PSMs are manifested in the following examples. (In these examples, pauses longer than three seconds are displayed with [… …].)
Episode 8:
Student J: The advantage of the Internet is that people ... can learn many things from it.
Retrospective comment: I was thinking about the tense of the verb.

Episode 9:
Student K: I think economy is more ... the best subject to be discussed in the talk shows.
Retrospective comment: I realized that I was not using the right adjective.

Episode 10:
Student L: It’s very good for ... very good to get useful information.
Retrospective comment: I realized that I should have changed the structure to convey my message.

Episode 11:
Student M: Most mothers tend to entertain their children by ... mothers accustom their children to sit in front of TV for hours to entertain them.
Retrospective comment: I couldn’t convey my original message and changed the structure to be able to convey it.

The above examples revealed that the nonfluent L2 speakers resorted to the grammatical-reduction subcategory of the resource-deficits PSMs, as well as to the error-repair, rephrasing-repair, and different-repair subcategories of the PSMs related to perceived deficiencies in their own output. In Episode 8, a nonfluent L2 speaker used the present tense of the verb and, thus, a simplified grammar. In the ninth Episode, a fluent participant made use of error-repair mechanism or made a self-initiated correction of an accidental lapse in his speech. The tenth Episode shows that a nonfluent L2 speaker repeated a slightly modified version of the adjective or employed rephrasing-repair PSMs. Finally, in Episode 11, a nonfluent participant had recourse to different-repair mechanism or changed his original speech plan and encoded his message differently. The results suggest that in contrast to the fluent L2 speakers, the nonfluent L2 speakers did not show facility at the application of the effective PSMs to overcome the halting effects of their linguistic and communicative deficiency sources. Besides, they apparently grappled with the frequently encountered problems in the monitoring phase of speech production and hesitated lengthily to handle the deficiencies in their own output. The output-related retrospective interviews further revealed that the nonfluent L2 speakers tended to abandon or replace their messages. The nonfluent
participants’ use of message abandonment and replacement mechanisms are illustrated in the following examples.

*Episode 12:*
Student N: Also, some viruses can be ... ... to your computer.
Retrospective comment: Here, I forgot the adjective detrimental and abandoned it.

*Episode 13:*
Student O: People should use both ... ... should use only good programs.
Retrospective comment: At first, I mistakenly wanted to say people should use both good and bad programs but then replaced it with good programs.

In sum, the above instances indicated that the nonfluent L2 speakers were not eager or simply were not capable of using effective PSMs and tended to abandon their messages or replaced it in a hesitant and nonautomatic way.

**Discussion**
This study drew on Leveilt’s (1989, 1999) psycholinguistic speech-production model to account for the cognitive processes involved in the fluent and nonfluent L2 speakers’ use of PSMs in L2 oral communication. The overall results indicated that the fluent L2 speakers employed useful and facilitative PSMs more often than the nonfluent participants in order to make their speech smooth and convey their intended messages. Specifically, they utilized a range of lexical PSMs such as approximation, use of all-purpose words, literal translation, circumlocution, and direct appeal for help more regularly than the nonfluent L2 speakers. This means that the fluent L2 speakers did not change their original macroplan and used different lexical PSMs to merely reformulate the preverbal plan and to compensate for their inadequate lexical resource-related deficits.

The fluent L2 speakers’ use of approximation strategy allowed them to omit or substitute specific specifications of the lexical chunk and to use another lexical item that had the same semantic features as the intended lexical item. Even sometimes, the fluent participants deleted many facets of the preverbal chunk so that they could only use a broad specification; that is, they used an all-purpose word or a general lexical item (e.g., thing and what-do-you-call-it) instead of their intended lexical item. In general, when the fluent L2 speakers made use of approximation and all-purpose-words PSMs, they used substitution strategy and modified or removed one or more conceptual specifications set in the preverbal message to compensate their insufficient knowledge of lexical items. The fluent L2
speakers, besides reforming the conceptual specifications of the lemma, also utilized L1 or L2 morphological and phonological encoding processes that led to the substitution plus strategy (Poulisse, 1993). The fluent L2 participants particularly benefited from the literal translation subcategory of the substitution plus strategy group (Dörnyei & Kormos, 1998) and unconsciously transferred a lexical item or a compound word from their L1 so that their speech seemed more automatic. The fluent L2 speakers’ adeptness at using these strategies made their speech more fluent and natural and helped them to bridge the communication gaps.

The fluent L2 speakers were also willing to reconceptualize or change the entire preverbal chunk by encoding the conceptual aspects of their planned lexical items distinctively (Poulisse, 1993). In other words, when they were not able to retrieve the proper lexical items, they circumlocuted or illustrated the features of them. The body of research on communication strategies has confirmed that the use of circumlocution strategy is useful in maintaining the flow of conversation (e.g., Dörnyei & Scott, 1995; Willems, 1987). In addition, when the fluent participants did not know a new lexical item, they sometimes tried to get direct help from the interlocutor by asking an explicit question; that is, they activated the interlocutor’s speech production mechanisms (Dörnyei & Kormos, 1998). To sum up, the L2 speakers’ employment of the aforementioned lexical PSMs assisted them to circumvent the difficulties they encountered in lemma retrieval. In fact, in contrast to the nonfluent L2 speakers, they did not abandon their messages and were proficient at maneuvering their available linguistic and communicative competence.

The results further indicated that the fluent L2 speakers were more sensitive to use a variety of stalling mechanisms and fillers than the nonfluent participant in order to avoid lengthy pauses. When the fluent L2 speakers experienced difficulty in any phases of speech production that needed more attention (i.e., during macro- and micro-planning, when the content and the form of the message are generated; while the preverbal plan is processed to generate the articulated message; in the monitoring phase; and during the comprehension of the interlocutor’s speech), instead of reducing or giving up their messages, they either utilized lexical PSMs or stalling mechanisms to sustain the conversation (Dörnyei & Kormos, 1998). They specifically used filling words and short phrases like well, you know, or actually and were skilful at temporizing and devoting more attention to speech processing in order to gain time to retrieve the necessary and appropriate lexical,
grammatical, or phonological items. Because of the serial nature of speech processing in L2 and inadequate processing time, the use of various kinds of stalling mechanisms and hesitation devices is the prerequisite of fluent speech production (Ellis, 1985; Kormos, 2006).

The fluent L2 speakers were also eager to use the meaning-negotiation strategies when they perceived problems in the speech comprehension system more often than the nonfluent participants. It means that when the fluent L2 speakers had problems in the interlocutor’s speech or simply could not hear the interlocutor, they tended to ask for repetition, clarification, or confirmation and also paraphrased or corrected the interlocutor’s speech (Dörnyei & Kormos, 1998). The use of meaning-negotiation strategies provided the fluent L2 speakers with more opportunities for comprehension and learning and improved the cognitive processes required for their L2 learning in that conscious production of speech results in internalizing linguistic forms (Long, 1996; Nakahama, Tyler, & Lier, 2001; Pica, 1996; Swain, 1995).

However, the output-related speech samples and the qualitative analyses of the retrospective interviews showed that, in search of a new lemma, the nonfluent L2 speakers abandoned or altered their original speech plan and did not show tendency to or simply were not able to maintain the macroplan and only adjusted the preverbal message (Dörnyei & Kormos, 1998). The findings related to the resource-related deficits PSMs further indicated that the nonfluent L2 speakers had recourse to the inefficient grammatical and phonological mechanisms. They usually transferred the grammatical and phonological rules from their L1 resulting in clumsy or awkward grammatical structures and pronunciations. In addition, they used simplified grammar that sometimes caused misunderstanding. This is largely due to the fact that the nonfluent participants had limited L2 linguistic and interactional competence and did not try to manipulate their linguistic knowledge or resort to the effective lexical PSMs to avoid the stumbling consequences of their deficiency sources.

Finally, the findings revealed that the nonfluent L2 speakers seemingly experienced frequent difficulties in the monitoring phase of language output both before and after articulation. This makes the speech of the nonfluent participants hesitant and unnatural because after detecting a problem, the conceptualizer received a signal, the encoding process came to a halt, and subsequently another
preverbal plan was produced (Levelt, 1989, 1999). Therefore, they had to stop many times and either self-corrected or self-repaired their speech in order to struggle to produce a new preverbal plan (Dörnyei & Kormos, 1998; Kormos, 2006). Specifically, the nonfluent participants resorted to error-repair, appropriacy-repair, different-repair, and rephrasing-repair mechanisms. When the nonfluent participants used error-repair mechanisms, they modified an accidental lapse that might have been occurred in any stage of speech production. After detecting the inappropriate or insufficient information, the nonfluent L2 speakers either used appropriacy-repair mechanisms to repair the utterance or employed different-repair mechanisms to encode different data. In other words, they modified both the macroplan and the preverbal plan partially or entirely. Moreover, the deficient L2 competence of the nonfluent L2 speakers led to their uncertainty and, thus, to using rephrasing-repair strategies (Dörnyei & Kormos, 1998). These findings point to the fact that the nonfluent L2 speakers did not make attempts to avoid their own-output deficiencies by utilizing useful lexical PSMs, stalling mechanisms, or meaning-negotiation strategies.

Conclusion

This study intended to examine the cognitive bases of L2 fluency from an integrative information-processing, speech-production management, and problem-solving standpoint. The overall results revealed that the fluent L2 speakers were adept at using facilitative and effective PSMs. They typically employed the PSMs related to the lexical-resource deficits, processing time pressure, and the perceived deficiencies in the interlocutor’s performance more regularly than the nonfluent L2 speakers to avoid communication impasses. In clear terms, the fluent L2 speakers retained the macroplan and merely modified the preverbal message to evade problems in lexical retrieval, employed a number of stalling mechanisms to devote more time and attentional resources to the serial nature of L2 speech processing, and utilized meaning-negotiation mechanisms to surmount difficulties in the speech comprehension phase. In contrast, the nonfluent L2 speakers were hindered by their linguistic and communicative deficiencies and mainly abandoned or altered their original speech plan. Moreover, although they confronted with frequent lapses in the lexical, grammatical, and phonological phases of speech production, they hesitated lengthily and did not show facility at employing efficient strategies such as lexical PSMs or stalling mechanisms to convey their messages and make their speech more natural. Instead, because the nonfluent L2 speakers suffered from deficient L2 linguistic and communicative resources, they resorted to
grammatical and phonological PSMs and mainly transferred the grammatical and phonological rules from their L1.

The examination of the cognitive processes underlying the speech-production management of differentially fluent L2 speakers in a comprehensive psycholinguistic framework is argued to have substantial implications for L2 instruction and assessment. A psycholinguistic approach illuminates the differences underpinning the information-processing mechanisms and the automatization processes of the fluent and nonfluent L2 speakers. Because the fluent L2 speakers did not change the macroplan and only modified the preverbal message, it can be assumed that the application of facilitative and time-gaining mechanisms counterbalanced or mitigated the stumbling effects of difficulties in lexical, grammatical, and phonological encoding phases. Moreover, the fluent L2 speakers’ high level of automaticity in using these mechanisms promoted the speech processing speed. The related literature has suggested that automaticity and proceduralization of L2 knowledge occurs in the formulating phase of speech production (de Jong & Perfetti, 2011; Towell et al., 1996); therefore, L2 teachers and curriculum developers should develop tasks that facilitate the formulating phase, and in particular the grammatical and phonological encoding mechanisms, of speech processing. In addition, L2 teachers should make L2 learners aware of a variety of effective PSMs such as lexical PSMs, fillers, and meaning-negotiation mechanisms and should integrate them into the language courses both implicitly and explicitly.

The findings have further implications for the L2 fluency assessment field. L2 research can integrate the neglected facets of L2 fluency as a cognitive information-processing perspective into its frequently studied performance-related aspects. The knowledge of the PSMs that the differentially fluent L2 speakers employed to surmount their possible linguistic and interactive deficiencies and their underlying cognitive processes can be used in educating well-prepared or knowledgeable raters and also in reshaping descriptors of fluency rating scales and, thus, in revising the existing merely output-related fluency rating scale.
Notes on Contributors:

Azizullah Mirzaei is Assistant Professor of Applied Linguistics at Shahrekord University, Iran. He has published or presented papers in (inter)national journals (e.g., TELL, IJAL, JTLS, Asia TEFL, English Language Assessment, Educational Psychology) or conferences (LLT, ILI, TELLSI, and KELTA). He co-authored a chapter on Assessing SL Pragmatics in The Cambridge Guide to SL Assessment. His research interests include: Sociocultural Theory, L2 Fluency, Interlanguage Pragmatics, Language Testing and Assessment, and Teacher Education.

Najmeh Heidari is an M.A. student of TEFL at Shahrekord University, Iran. She has presented or published papers in (inter)national conferences (e.g., TELLSI & ILI) and journals (e.g., Asia TEFL, Educational Psychology). Her areas of interest include: A Socio-cognitive Approach to L2 Fluency, Communication Strategies, Multiple Intelligences, and Language Learning Skills.

References


**Appendix**

**Problem-Solving Mechanisms (PSMs) Instrument**

University: ....... Major: ....... Semester: .......

Age: ...

Gender: .......

How often did you have real-life face-to-face communication during high school, B.A. studies, and M.A. studies?
Always ☐ Often ☐ Sometimes ☐ Rarely ☐ Never ☐

Have you ever been to an English-speaking country?
No □ Yes □ If yes, how much time did you spend there? ..........

**Instruction:** Please read the following items, choose a response indicating how often you employ each solution given to tackle your speaking-related problems while communicating, and write it in the space provided after each item.

1. **Never** 2. **Sometimes** 3. **Generally** 4. **Always**

1. I reduce the message or leave it unfinished by avoiding certain language structures or topics when I encounter some language difficulty while communicating in L2.
2. I substitute the original message with a new one when I am unable to retrieve the appropriate words to communicate my intended message.
3. I use a single alternative item (e.g., a superordinate or a related term, such as using *mouth* instead of *beak*) that shares semantic features with the target word or structures when I forget the appropriate L2 lexical items.
4. I create a nonexisting L2 word by applying a supposed L2 rule to an existing L2 word or by compounding words when I have difficulty saying a word.
5. I translate literally a lexical item, an idiom, a compound word, or a structure from L1 or L3 to L2 when I have difficulty saying the appropriate L2 word.
6. When I am unable to remember the appropriate L2 words, I abandon the execution of a verbal plan and leave the utterance unfinished and communicate the intended message according to an alternative plan.
7. I exemplify, illustrate, or describe the properties of the target object or action when I have difficulty remembering the exact L2 word.
8. I try to elicit help from the interlocutor either directly by asking an explicit question or indirectly by expressing lack of a needed L2 item.
9. I change the meaning and syntax of a lexical entry through transfer or overgeneralization when I encounter deficiencies in L2 grammatical knowledge.
10. When I have deficiencies in L2 grammatical knowledge, I use simplified grammar in the belief that the interlocutor will be able reconstruct the grammatical meaning from the context.
11. In an attempt to retrieve and articulate a lexical item, I say a series of incomplete or wrong forms or structures (e.g., *it's kind of*) before reaching the optimal form.
12. I use nonlexicalized filled pauses (er, uh, mhm) or filling words and gambits (e.g., you know, actually) to keep the conversation going.
13. I repeat my own utterances or the interlocutor’s utterances in order to gain time and keep the conversation going.
14. I make self-initiated corrections and correct inappropriate or inadequate information in my speech when I realize that I was not using the right word or structure.
15. I change my primary message or repeat the slightly modified version of a word or phrase when I recognize inappropriate words or structures in my speech.
16. I ask questions to check whether the interlocutor understands my message or to check the correctness of my speech.
17. I request repetition or explanation when I have difficulty hearing or understanding something properly in L2 communication.
18. I express my lexical, grammatical, and conceptual misunderstandings either verbally or nonverbally when communicating in L2.
19. I request confirmation or paraphrase the interlocutor’s message to ensure that what I heard or understand is correct.
20. I guess the intended meaning of the interlocutor based on contextual discourse clues.