

Speech Processing and Fluency Model: New Directions for improving Speaking Skills

Rakesh Ramachandran, J Sundarsingh

Abstract The framework of complexity, accuracy and fluency (CAF) has gained currency in the area of applied linguistics research, language teaching and testing in order to analyze the oral language performance of second language speakers ever since Skehan (1996) proposed it. Many studies have tried to relate CAF framework with the various models of speech processing in order to ground it psycholinguistically. However, there is a need to offer more clarity regarding the psycholinguistics of speech processing and its specific relationship with the complexity, accuracy and fluency. This paper is an attempt to situate CAF framework on Levelt's Model of Speech Processing (Levelt 1989) thereby providing it a stronger psycholinguistics foundation. It also aims to provide a more specific explanation about what gets reflected as accuracy, complexity and fluency in speech. The paper concludes by offering more specific questions that second language teachers in particular and applied linguists in general may consider exploring the concept further.

Key words: *Speech processing, conceptualization, formulation, articulation, coordination, psycholinguistics, oral performance.*

I. INTRODUCTION

Speaking is a complex phenomenon not just because it uses sound symbols of language arranged on the basis of abstract rules of syntax for communicating the intention of the speaker, but also because it is online and delivered in real time taking into considerations the sociolinguistic relationships the speaker has with the listener(s). According to Bygate (2006), speaking is more complex because it is one off unlike writing and the editing/correction/repair involved in speech is public. Speaking in the second language can be even more complex since many of the second language speakers may not have sufficient and appropriate linguistic resources to encode the message and the processing skills to deliver it fluently. Moreover, affective factors which include lack of confidence, introversion etc can also pose challenges (cf Krashen 1983) to some of the second language speakers.

When these complexities make speaking a difficult language skill, teaching speaking becomes even more challenging for the language teacher. The teacher will have to parallelly develop the linguistic resources and the speed of processes involved in encoding the message in addition to minimizing the effect of debilitating affective factors. Therefore, a thorough understanding of the various processes involved in speaking as well as the diverse resources required for encoding the message linguistically will be helpful for the language practitioners who aim to develop the speaking skills of the learners. In addition to the classroom practitioners, the material producers, syllabus designers, language testers and

researchers also need to have an in depth knowledge of various processes and resources involved in speaking.

1. Speech Processing – An Overview

The literature on speech processing is replete with various theories and models of speech processing. Some of the models include Dell (1986), Levelt (1989), Van de Merve (2001) etc. Even though these models vary on certain aspects of speech processing like the option of giving feedback to higher level processes, they seem to converge generally on the various processes involved in processing the speech as pointed out by Schriefers and Vigliocco (2001). Levelt (ibid), Van de Merve (ibid), Kormos (2006) and Bygate 2006, 2011) agree on the notion that speech is processed in four stages – Conceptualization, Formulation, Articulation and Monitoring. Similarly, there is also general agreement on the notion that each of these stages except Monitoring depend on certain knowledge stores for linguistic resources during the act of speech processing. There also seems to be consensus on the idea that speech is processed incrementally (in units/parts and not as whole sentences) and that the length of these units becomes smaller as they move from Conceptualization to Articulation (bygate 2009 b). In short, there is general agreement between various models and scholars regarding the stages involved in speech processing, presence of linguistic knowledge stores and incrementality during speech processing. Since the central aim of the present paper is to synthesize the notions of speech processing in various models to offer a comprehensive view of speech processing, the next section will focus on each of the processes involved in speech processing.

1.1. Conceptualization

According to Levelt (ibid) Conceptualization is the first of the series of processes involved in speech production in which the intention of the speaker is transformed into a 'Preverbal Message' which functions as an input for the next process. Even though this looks very simple and straight forward, the conceptualization process is really complex because of the sub-processes and the cognitive load involved in it. Figure 1 provides an overview of the various sub-processes involved in generating a 'Preverbal Message' based on the communicative intention of the speaker.

Revised Version Manuscript Received on 10, September 2019.

Rakesh Ramachandran

J Sundarsingh, (email: sundarsingh@karunya.edu).

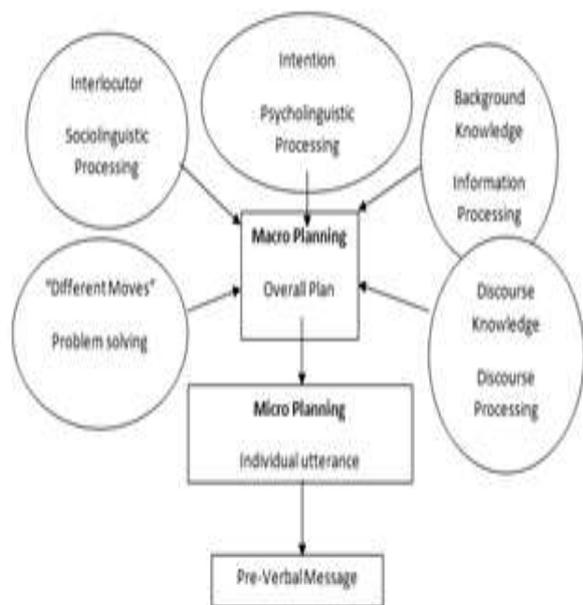


Figure 1
Conceptualization Process

The figure reveals that the communicative intention of the speaker is first converted into an ‘Overall Plan’ through ‘Macroplanning’ of the whole discourse which may be a speech, conversation, debate etc. In short, Overall Plan is a product of the Macroplanning process and this overall plan contains ideas/information, their sequence and various strategic moves required to achieve the communicative purpose. ‘Macroplanning’ of speech takes into account various aspects like the relationship between the speaker and the listener(s)/interlocutor(s), the context of speech and the type of discourse involved before the ‘Overall Plan’ of the speech is generated. In order to figure out the relationship between the speaker and the listener in the linguistic context of the speech, sociolinguistic processing of the situation is required. Similarly, Discourse processing and information processing are required to select the appropriate discourse and the relevant information to be included in the speech respectively. In addition to that, aspects of ‘Problem Solving’ are involved in converting the communicative intention into a speech act. Some amount of psycholinguistic processes will be involved before the speaker decides upon the intention to be communicated through the whole speech. This means that the sub-process of Macroplanning itself are a very complex process and place a great demand on the attentional resources of the speaker’s working memory.

Once the ‘Overall Plan’ of speech is generated, it is stored in the Working Memory and based on it, individual utterances that are consistent with the plan is generated through the sub-process of ‘Microplanning’ (Bygate 2009). Figure 1 also shows that the output of ‘Microplanning’ is a ‘Pre-verbal Message’ of the individual utterance. This preverbal message is also stored in the Working Memory for the subsequent processes to act to generate the final output for the listener. It is this ‘Preverbal Message’ that acts as the input for the immediate next level process which is ‘Formulation.’

1.1. Formulation

Formulation is the second in sequence of the four

processes involved in the processing of speech. It is through the formulation process that the ‘Pre-verbal Message’ which is non-linguistic is encoded linguistically. In order to convert the message into a linguistic form, the ‘Preverbal Message’ is first lexically encoded. Lexical encoding is a sub-process of ‘Formulation’ in which the appropriate lemmas stored in the lexical store is activated and selected in order to encode the message verbally (Kormos 2006). These lemmas then will be aligned in such a way as to form a grammatically correct unit. According to Schriefers and Viglioco (ibid), the activated and selected lemmas first undergo functional level processing in which the lemmas are sequenced based on the relationship between the participants (subject, object etc) of the ‘Preverbal Message.’ Then it undergoes positional level processing in which they are linearized based on the syntactic rules of the language to be spoken. The functional and positional level processing of lemmas together form the grammatical/syntactical processing which is a sub-process of ‘Formulation.’ In order to process the lemmas syntactically, the syntactic patterns and structures stored in the syntactical store will be used (see figure 2).

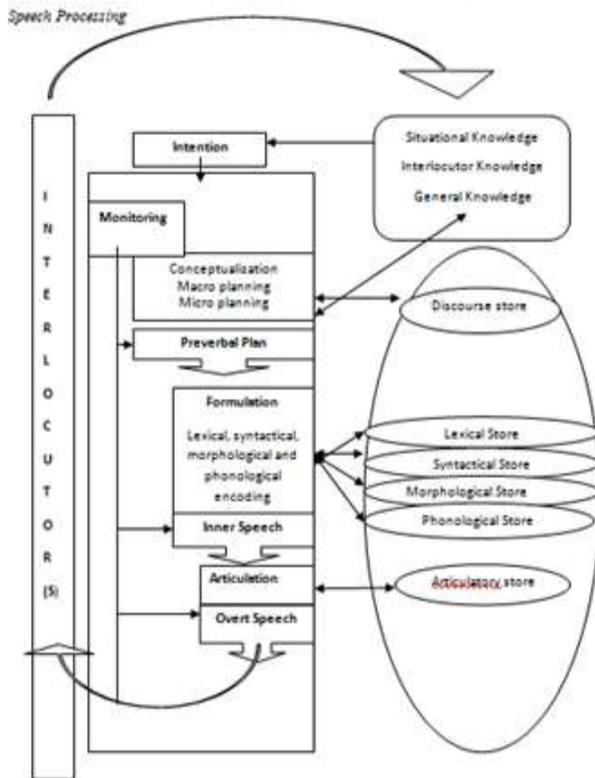
After encoding the message grammatically, it is morphologically encoded by selecting appropriate morphemes from the morphological store. This sub-process is called morphological processing and the output of the morphological process is then sent for phonological processes. During phonological processing, the phonemes corresponding to the morphemes of the message are selected from the phonological store along with metrical patterns of stress and rhythm that suit the message. After the selection of required phonemes and metrical patterns, the whole message is syllabified based on the syllable rules of the language in question. At this stage, an inner speech is generated which has specific information regarding the syntactic structures, morphemes and syllables and stress patterns of the particular language. In short, the output of the formulation process is an ‘Inner Speech’ which may be sent to the next level process called the articulation process. However, it is equally important to note that this inner speech is monitored and revised if it does not match the Preverbal Message

1.2. Results & Discussions

The next phase of speech processing is articulation in which the ‘Inner Speech’ is converted into ‘Overt Speech.’ During this process the ‘Inner Speech’ is transformed into actual overt utterance based on the information stored in the articulatory store which specifies the muscular movements of the organs of articulation (tongue, lips, jaws, glottis, vocal cords etc). Bygate (2009 b) points out that speaker can monitor the ‘Overt Speech’ and revise, edit, repair or modify it if it does not suit his communicative intentions. However, all these modifications will become public and can be heard by the listener/interlocutor unlike the modifications in the previous stages.



Figure: 2 speech processing



1.1. Monitoring

The picture of speech processing that emerges from the discussions in the previous section makes clear that speech is processed in three stages and that each of these three stages draw on its corresponding knowledge store(s) in order to linguistically encode the communicative intention of the speaker. There is also a fourth stage called monitoring which acts between and at the end of all these processes in order to verify the correctness and suitability of the product of each of these processes. Another point that emerges from the discussions is that the higher levels coordinate the lower levels and the lower levels act upon the frames offered by the higher levels with the 'Conceptualization' offering the overarching frame or 'Overall Plan' of the entire speech (Bygate 2006). However, it should be noted that the lower levels can provide feedback to the higher levels when necessary. Sometimes the lower levels may not have sufficient resources to map language to the frame offered by the higher level. In such cases, the lower levels give feedback to higher levels and the higher levels may alter the frame in such a way as to suit the resources available.

While all these four processes work parallelly, one should keep in mind that the higher levels move slightly ahead of the lower levels in order to offer an uninterrupted flow of speech. This means that the higher levels will be processing the subsequent part of the sentences while the lower levels will still be processing the knowledge stores for resources to encode the frame provided by the higher levels.

II. SPEECH PROCESSING AND CAF

The CAF (Complexity, Accuracy and Fluency) framework proposed by Skehan (1996) to analyze linguistic performance is widely used in applied linguistic research and practice. This framework suggests that both oral and written

performance can be analyzed based on how complex, accurate and fluent they are (Bygate 1996 & 2001, Ellis 2005 & 2009, Lynch and Mc Lean 2001). A reasonably good oral performance will be characterized by diversity of vocabulary and structures (complexity), correctness of syntax, morphology and pronunciation (accuracy) and flow of linguistic elements which do not cause stress for the listener (fluency). In other words, any linguistic performance which is not complex, accurate and/or fluent cannot be termed as ideal. Therefore, CAF framework is widely used in language testing to rate the performance of the test takers (Mehnert 1999, Ortega 1999, Tajma 2003).

However, CAF framework is strongly rooted in the psycholinguistics of speech processing as elucidated by the Levelt's model (Bygate 2009 a). Complexity of a performance could be the result of the richness of resources a language user has in his various stores (discourse, lexical, syntactical, morphological, phonological and articulatory). If a speaker has a great deal of items in all these stores, he will be able to select a range of appropriate items from them to express the same meaning so that it reflects as complexity of expression. Similarly, the ability of the speaker to correctly select the required item from these stores and join them together based on the syntactical rules of the language will be reflected as accuracy of linguistic performance. Furthermore, the speed with which a language user processes the stores for resources and the degree of automaticity achieved in encoding processes will be reflected as fluency of linguistic performance.

Form the above conceptualization of CAF framework, one should be able to understand that as a person advances in his/her linguistic abilities, the performance tends to be more complex, accurate and fluent. Psycholinguistically, this means that the person increases the resources in his stores (complexity development), enhances his/her ability to correctly use the resources for encoding the message (accuracy development) and speeds up his/her ability to access the resources by developing automaticity (fluency development) through repeated practice or use of language. Therefore, development of linguistic capacities can be explained using the Levelt's model.

III. CONCLUSION

It becomes clear from the previous discussions that the performance features of oral language output could be grounded on the psycholinguistics of speech processing based on the model proposed by Levelt. This offers more clarity to conceptualize the complex skill of speaking from a comprehensive perspective. It also offers the language teaching professionals in general and the professionals who are focused in the area of speaking in particular to conceptualize speaking based on different processes and knowledge stores involved. Such an analytical perspective may pose new questions for developing speaking skills.

Rather than asking how to develop speaking skills, the language teacher interested in the development of the speaking skill should find means to populate the



lexico-morpho-syntactic store of the learner so that the spoken output of the learner would be more complex. The teacher should also enable the learner to select the correct and appropriate lexis, structure, morpheme, phoneme and intonation so that the spoken output becomes more accurate, besides developing the speed of processing of the learners which would reflect as gains in fluency. Similarly, other language professionals which include curriculum developers, material developers, language testers also have more specific and focused questions to ask themselves before embarking on their task, if they attempt conceptualized speaking in line with the perspective offered in this paper.

performance. *Studies in Second Language Acquisition*. 21: 109-148.

28. Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford. Oxford University Press.
29. Tajima, M. (2003). *The effects of planning on oral performance of Japanese as a foreign language*. Unpublished Dissertation, Prude University.

REFERENCES

1. Bygate, M. (1996). Effects of task repetition: appraising the developing language of learners.
2. In Willis, Jane. & Willis, Dave. *Challenge: Change in Language Teaching* (pp. 136-146). London: Heinemann.
3. Bygate, Martin (2001). Effect of task repetition on the structure and control of oral language. In
4. M. Bygate, P. Skehan & M. Swain (Eds), *Researching pedagogic tasks: Second language*
5. *Learning, Teaching and Testing*. (pp. 23-48). Harlow: Longman.
6. Bygate, M. (2005). Oral second language abilities as expertise. In K. Johnson (Eds), *Expertise in*
7. *Second Language Learning and Teaching*. (pp. 104-127). London: Palgrave Macmillan.
8. Bygate, M. (2006). Areas of research that influence L2 speaking instruction. In Euso-Juan and A.
9. Martinez-Flor (Eds). *Current Trends in the Development and Teaching of the Four Language Skills* (pp. 159-186). Berlin: Walter de Gruyter GmbH & Co.
10. Bygate, M. (2009 a). Teaching and testing speaking. In Michael H. Long and Catherine J.
11. Doughty (Eds). *The Handbook of Language Teaching*. (pp. 412-440), Wiley-Blackwell, Chichester.
12. Bygate, M. (2009 b). Teaching the spoken foreign language. In: Knapp, Karlfried, Seidlhofer,
13. Barbara (Eds). *Handbook of Foreign Language Communication and Learning*.
14. *Handbooks of Applied Linguistics* (6). (pp. 401-438). Berlin: De Gruyter Mouton.
15. Ellis, R. (2005). Planning and task based performance: Theory and research. In Ellis, R. (Eds).
16. *Planning and Task Performance in a Second Language*. (pp.3-36), Amsterdam: John
17. Benjamins Publishing Company.
18. Ellis, R. (2009). Differential effects of three types of task planning on the fluency, complexity
19. and accuracy in L2 oral production. *Applied Linguistics*. 30/4: 474-509.
20. Kormos, J. (2006). *Speech Production and Second Language Acquisition*. London. Lawrence
21. Erlbaum Associates, Publishers.
22. Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, Mass: MIT Press
23. Lynch, Tony and Maclean, Joan. (2001). 'A case of exercising': effect of immediate task
24. repetition on learner's performance. In Bygate, Skehan, and Swain, (eds). *Researching Pedagogic Tasks and Second Language Learning, Teaching and Testing*. (pp. 141-162). Harlow. Pearson Education.
25. Mehnert, Uta. (1998). The effects of different lengths of time for planning on second language
26. performance. *Studies in Second Language Acquisition* 20 (1): 83-108.
27. Ortega, L. (1999). Planning and focus on form in L2 oral