Speech as a Psycholinguistic Process: The Missing Link in Oral Lessons

Christine C. M. Goh

Introduction

Speaking appears to be an easy task for some people. Words flow seamlessly as ideas are conveyed through utterances with flawless grammar. Unseen to listeners, however, are myriads of complex mental processes taking place. For fluent speakers, these processes work interactively, often automatically, and in harmony, making optimal use of the limited processing capacity of the working memory. Others, on the other hand, may experience difficulties at various levels of processing, which may be further complicated by social and linguistic issues.

This article introduces some theoretical perspectives on speech as a psycholinguistic process that take into consideration both structural and functional aspects of language. The final section will discuss implications for two common misconceptions about learner behaviour and classroom practice with regard to the teaching of oral skills.

Speech as process, particularly a psycholinguistic one, is one of the least discussed aspects of speaking. Consequently, its implications for classroom teaching are often overlooked. Teachers are familiar with speech as a “skill” and speech as a “product” because these aspects are emphasized in text books, language curricula and examinations. Examples abound in the English Language Syllabus 2001 published by the Ministry of Education, Singapore. Learning outcomes for both primary and secondary schools include such skills as retelling a story, describing a picture, expressing thanks and good wishes, explaining why and how something happens, giving instructions, explaining own views, initiating and sustaining conversations, and supporting opinions with reasons. Products of speech
are the various spoken text types or speech genres, which include presentations, narratives, conversations, debates, interviews and oral summaries. Speech as a product is also characterized by specific features such as discourse structure, spoken grammar (McCarthy and Carter, 1995; McCarthy, 1998), as well as phonological features, such as pronunciation, rhythm and intonation.

Considering speech as skill and product can help teachers identify lesson objectives and select learning tasks for developing learners' speaking ability. Moreover, being observable learning behaviours and outcomes, sub-skills and oral production can be readily assessed. These considerations alone, however, are insufficient for a comprehensive approach to teaching speaking in the language classroom. This is because the sub-skills and products which we hope our students can demonstrate are the result of interplay of psychological processes and other factors. This article seeks to review some current perspectives on the speech process by answering three key questions:

- What is the nature of speech processing?
- What factors influence the speed of processing?
- What are the implications of the above for the language classroom?

**Review of Research**

**The Nature of Speech Processing**

The processes of language production can be divided into those that create the skeleton of an utterance and those that flesh the skeleton out (Bock and Levelt, 1994).

The term "speech" is used here to refer to any kind of oral language production ranging from that which is spontaneous, exploratory and ephemeral to the type that is drafted, edited and rehearsed (Willis, 1996). In fluent conversations, a speaker may generate two or three words per second by retrieving them from a memory store that contains tens of thousands of items (Levelt, Roelofs and Meyer, 1999). This seemingly effortless performance actually involves underlying processes that are exceedingly complex. Two areas that researchers have focused on are the nature of speech processing and factors that affect the speed of processing. Language learning researchers have also examined the effects of pre-task planning on accuracy, fluency and language complexity of oral language performance.
Processes in Speech Production

One framework for analysing the core processes is proposed by Levelt (Levelt, 1989; Levelt et al., 1999) which can be broadly classified as conceptual preparation, formulation and articulation.

Conceptual preparation refers to how speakers select the topic/content, or the notion/information to be expressed. If a topic has already been established, speakers have still to select relevant ideas which can reflect their message and communicative intention. Message selection will depend on speakers' background or “encyclopaedic” knowledge about the world. The message to be expressed may exist only vaguely, as a mental model or simple meaning components (a “gist”) which have to be mapped on to specific words in the speakers' mental lexicon (Garman, 1990).

Formulation is the process by which the concepts selected are converted into language. Special considerations are paid to word choice and grammar. The concepts are “fleshed out” through a stringing together of words in the right order (syntax). Markers of tense, mood, number etc. have also to be indicated by selecting appropriate bound morphemes (e.g. -ed, -s, -ing). The formulated utterances are then conveyed through the activation and control of specific muscle groups of the articulatory system (consisting of vocal tract, larynx and lungs), which carries the linguistic message to the listener in the form of sound waves. This is the process of articulation.

These three processes of conceptual preparation, formulation and articulation are often subsumed under the term “planning”. The amount of time taken for planning will depend on the type of speech to be produced. It will depend also on the context of use in which it is produced and the linguistic competence of the speaker. For example, in spontaneous interactive talk such as conversations, speakers have very little time to plan. Conceptualizing, formulating and articulating all take place at much the same time. This is when real-time speech problems such as dysfluencies are most noticeable. In contrast, longer planning time is usually available for more formal oral production such as debates and presentations. The result is that the outcome for the first may be 'messier' when compared with the latter, which is edited and rehearsed. The potential core levels or components of planning are the same nevertheless. While there is general agreement about the processes involved, there is still uncertainty about how these processes take place in real time. The main question is whether they take place sequentially or simultaneously.
Serialist and Connectionist Models of Processing

In the late 1950s and early 1960s information processing was presented as a number of cognitive stages which follow one another in a neat and orderly manner. This serial or linear model was severely criticized in the 1970s by researchers who derived new evidence about the interactive nature of processing, involving both top-down (knowledge/schema-driven) processes and bottom-up (data/text-driven) ones. While some experimental testing has provided new empirical support for sequential processing (Schriefers, Meyer and Levelt, 1990), the view that has gained currency is that cognitive processes are interactive and occur simultaneously through a process of spreading activation of interconnected neural networks in the brain (Dell, 1986; Rumelhart, McClelland and the PDP Research Group, 1986; Bechtel and Abrahamsen, 1991). One limitation of much of the work in computational modelling is that it focuses mainly on the word or the sentence levels, and not at the level of discourse where most speech production typically occurs. Although there is much more to be known about how the mind processes language for speech, the information that is already available still has implications for teachers and researchers involved in the areas of speaking and listening.

The Role of Monitoring

A discussion of speech processing is not complete if it focuses only on mental processes that act directly for the generation of speech. Metacognitive processes which manage and regulate these cognitive processes are equally important and are an integral part of the process of speech production. In general, every individual is able to monitor their overt speech output. We notice errors in pronunciation, grammar, dysfluencies and other problems commonly associated with speech production (Levelt, et al., 1999). This is done both during and after speech production. (Some writers prefer to use the term “evaluation” when the task has been completed.) For various reasons, however, this is not always done. Speakers may also choose to ignore errors rather than make “self-repairs” to improve what they have just said.

Besides linguistic demands, speakers also have to balance pragmatic demands and be sensitive to “reciprocity conditions” (Bygate, 1988). Speaking is a reciprocal or two-way activity. In most situations, speakers are in direct contact with their listeners. This proximity allows speakers to obtain both explicit and implicit feedback on what they are saying. In other situations, such as speaking to a distant
audi ence via radio or television broadcasts, speakers will need to anticipate their listeners' reactions. Basic pragmatic demands include speakers' assessment of the power relationship between themselves and the listeners as well as the context in which speech is produced. In situations when speech is delivered in face-to-face interactions, monitoring is done in real-time. This may put further pressures on the cognitive processes already in operation. On the other hand, when speakers have the benefit of time and distance, monitoring may be done during preparation and rehearsal. Communicatively competent speakers not only produce utterances that are accepted by their listeners in terms of grammar, vocabulary and pronunciation, they also evaluate the appropriateness of what they say in the light of these reciprocity conditions.

**Factors that Influence the Speed of Processing**

A second area that researchers focus on is the factors that influence the speed of processing, which is reflected in the fluency of a person's speech. Specific features such as appropriate word choice, grammaticality, relevant and interesting content are other indications of how effectively speech is processed. Bygate (1998) outlined several matters that can influence the processes we have just described. We shall now discuss these and other issues in the light of how they can speed up processing, or alternatively, put constraints on it.

1. **Discourse routines**: In conversations or other exchanges, speakers draw on conventionally acceptable phrases and procedures, or "routines", to initiate or end talk, negotiate or clarify meaning, take, keep or pass on a turn and manage the agenda. These routines are the result of explicit and implicit language socialization, through both formal and informal interactions an individual has experienced. For every exchange, speakers have to identify a message structure that is appropriate for the context. They also have to employ the relevant skills for specific parts of the routine. A speaker who is aware of but unfamiliar with these conceptual demands will need to spend time accessing such information from the memory store.

2. **Lexical access**: Speakers' ability to recall and select from words stored in long-term memory is another factor that influences fluency (Levelt, 1989; Levelt et al., 1999). The access may be instantaneous or it may require several intermediate steps before the final desired word is recalled. At the same time, a speaker has to balance two classic aspects of language selection: syntagmatic and paradigmatic alternatives (Bygate, 1998). In the former, the speaker has to
choose words that can be strung together in logical and grammatical sequences. In the latter, the speaker needs to decide the “best” word to use out of a number of possible alternatives (e.g. whether to use pretty, beautiful or attractive). Speakers who have a larger store of vocabulary will thus potentially have more alternatives to choose from, but this does not necessarily enable quicker processing. Some speakers lose processing time when deliberating which word to use, and may even lose their turn (Hughes, 2002). Using pre-fabricated “chunks” of language can also speed up processing. Chunks are formulaic phrases for specific occasions, for example, Congratulations on your promotion, pardon my ignorance, and collocational units (words that usually go together), for example, a tall order, a terrible mistake, nook and cranny, etc.

3. **Automatization:** Cognitive processes that have been well learned become automatized through constant use and rehearsal (Shiffrin and Schneider, 1977). They make little or no demand on processing capacity because they do not require attention. The working memory is therefore free from one level of processing to attend more closely to the others. At the conceptualization phase, certain types of information or routines may be automatized because of the speaker’s prior knowledge of facts, social or academic conventions and specific text structure. Formulation processes that can become automatic include word selection, grammar rules and pronunciation. Speakers may experience different degrees of automatization in different situations and according to differing demands made on the output. For example, when talking about a familiar topic during conversations with friends, a person might process his or her speech quickly and efficiently. The same topic, however, may require greater attention and controlled planning when it is communicated to a large audience in a formal situation.

4. **Facilitation features:** In spontaneous speech, speakers make use of a number of facilitation features to help them cope with limited processing capacity and time pressure (Bygate, 1987). They may use more co-ordination of simple clauses (e.g. joining short utterances with and, or, because) instead of attempting to produce long complex sentences with many clausal embeddings (e.g. combining main and subordinate clauses that make use of whereas, when, if). Grammar in spoken language is also simplified when speakers resort to ellipsis or omitting parts of a sentence so as to speak economically. The parts that have been omitted are contents or ideas that are readily inferred from context and the co-text. For example: got to go, over there, if not. They also use many fixed or formulaic expressions as fillers, for example, What you’ve said is
extremely interesting. These facilitation strategies are important for “buying processing time” (Hughes, 2002).

The review so far discussed the psycholinguistic processes and conditions for effective processing. In the next part of the review, we will focus on some research which has specific relevance to language learning.

**The Effects of Pre-task Planning in Language Learning**

Language learners experience a great deal of stress when they have to speak spontaneously. When this happens, their attention will typically be directed to conveying their meaning, which is often at the expense of accurate forms. The absence of time pressure, on the other hand, might allow learners to consider how to incorporate appropriate but difficult forms. This has important implications for language learning and several studies have sought to investigate whether allowing language learners time to plan can significantly enhance their speech production.

In a sense, it seems like a foregone conclusion that planning will improve talk. Or does it? Does pre-task planning time improve all aspects of speech? Most studies have focused on the formal aspects of language output such as fluency, sentence complexity, vocabulary range and grammatical accuracy. In other words, the focus was predominantly on the level of formulation. The results show that pre-task planning can give learners the opportunity to pay attention to those “problem” areas of language which they have yet to master (or “automatize”) (Ellis, 1987; Skehan, 1998). It encourages them to make explicit function-form relationships (i.e. appropriate grammatical features for a specific type of communicative task). In addition, by experimenting with a greater variety of forms, learners also increase language complexity in their oral performance (Crookes, 1989), as well as improve their chances of internalizing these forms (Ellis, 1987).

In another similar study, Ortega (1999) also concluded that pre-task planning could improve language complexity and speaker fluency. Findings about grammatical accuracy were, however, inconclusive. She argued that pre-task planning may be more beneficial for higher proficiency learners. Low proficiency learners, on the other hand, have limited linguistic resources to work with, so extra planning time may not be very useful. One interesting finding is that that learners attended to form both before and during the oral task. In other words, the learners were not only making decisions about structure during the planning stages, they
also monitored themselves when they spoke. A small percentage of the subjects reported that they did not think pre-task planning time was helpful.

**Conclusion**

The apparent ease with which speech is produced belies the complex cognitive processes involved and the factors which influence it. Very often, these processes are overlooked in the classroom where attention is invariably on the final product. It is hoped that by drawing teachers' attention to speech as a psycholinguistic process, they can adopt a more balanced approach when planning lessons on speaking and listening. These can facilitate better oral performance among language learners which can have a cumulative and lasting impact on their language development as a whole.

**Speech as a Psycholinguistic Process Implications**

Below is a discussion of how knowledge about speech as a psycholinguistic process may help modify two common misconceptions related to the teaching of oral skills:

**Misconception 1: My pupils don’t speak up because they are “shy”**

Not all pupils who keep quiet in their language classes are necessarily shy by nature. Some of them may want to speak up but they are unable to cope with the cognitive demands of one or all of the core processes:

- **Conceptual preparation** (*Don’t know what to say*): This could be due to insufficient background or content knowledge, or the inability to select something that is appropriate for the task.
- **Formulation** (*Don’t know how to put things in a “proper sentence”*): The pupil may have a notion of what to say, but experiences difficulty in translating that mental model into more precise language, or selecting the right word, or may not know specific key words in the target language.
- **Articulation** (*Don’t know how to pronounce specific words in the target language*): The pupil has formulated a proposition, but may not be able to articulate it clearly. In some cases pupils may be genuinely embarrassed by their poor pronunciation.
It is worth remembering that learners may not experience the same kind of difficulties all the time. These difficulties depend greatly on the nature of the speaking task and the pupils’ perceptions about different matters, for example, the degree of formality, time pressure, interest level, background knowledge, etc.

**Misconception 2: Pre-task planning time always leads to better oral performance**

This may be true only to a certain extent. There are, however, at least two reasons why giving learners time to prepare what they have to say will not necessarily result in greater accuracy in their oral production.

(i) First, giving learners some time to plan their summaries, presentations, and stories may allow them to improve the content, but it does not always lead to better formed sentences, greater language complexity or better delivery. Left on their own to use the planning time, some pupils may not focus on form (i.e. consider the relevant grammatical aspects necessary for the communicative task at hand). They may also ignore the need to rehearse aloud what they want to say or to check up the pronunciation of key words. In other words, unstructured or unguided planning may have little significant impact on crucial aspects of language development, such as greater accuracy in grammar and pronunciation.

(ii) Second, without teacher intervention, weaker pupils may hit the “proficiency ceiling”. Pupils may be instructed to “pay attention to your grammar and pronunciation”, but the weaker ones have limited linguistic knowledge to fall back on when planning and monitoring their eventual output. If left unsupported, they may not know what formal features are appropriate for a specific oral task and will not be able to spot their mistakes. In other words, they will continue to make the same mistakes even with the benefit of more time and cognitive space to plan what they have to say. On the other hand, the stronger pupils, not having to work under the pressure of time, may be able to draw on their declarative knowledge about specific aspects of language and apply it in the task.

(iii) Third, this notion of accuracy in form should not be confined to isolated grammar items alone. Formal accuracy should include macro
aspects such as text type-related grammar and discourse routines/conventions. While this connection may be immediately clear in the context of teaching writing (i.e. compositions), it is often overlooked for teaching listening and speaking. Oral tasks should therefore include the production of spoken text types that are relevant to the three areas of language use in the EL Syllabus 2001: information, interaction and literary expression/response.

References


