Monitoring in Speech Production and Second Language Acquisition

การตรวจสอบในการผลิตคำพูดและการรับรู้ภาษาที่สอง

Jan Stevener[1]

Abstract

This paper aims to understand the cognitive process of monitoring which is considered crucial to language production and second language acquisition. Levelt (1989), incorporating previous work from Krashen (1982) and others, made the monitor an essential part of his speech production model. His concept of the monitor saw its basic function in checking the pre-verbal and verbal language output; in case a speech error is detected the monitor would immediately stop language production processes and go back to a previous step in the language production process. Other influential models also assume monitors (MacKay, 1987; Levelt, 1989; Schade, 1999). Newer research, however, shows that the role of monitoring processes is far more complex and that in learners with more advanced language skills and a higher degree of automaticity the monitor will make more covert repairs.

Keywords: Monitor; Monitoring; Krashen; Levelt; Second Language Acquisition; Speech Production

[1]Lecturer from the German Section, Department of Western Languages, Faculty of Humanities, Ramkhamhaeng University.

Email: janstevener@yahoo.de

บทคัดย่อ

บทความชิ้นนี้มุ่งเน้นศึกษากระบวนการทางปัญญาในการตรวจสอบ ซึ่งถือว่ามีความสำคัญอย่างยิ่งในการผลิตภาษาและการรับรู้ภาษา Levelt (1989) ซึ่งรับแนวความคิดที่ปรากฏมาก่อนของ Krashen (1982) และของนักวิจัยคนอื่น ๆ ได้ยกให้การตรวจสอบ (monitor) เป็นส่วนสำคัญส่วนหนึ่งของทฤษฎีการผลิตคำพูดของเขา หน้าที่พื้นฐานของการตรวจสอบตามแนวความคิดของ Levelt ก็คือ การตรวจสอบภาษาก่อนและระหว่างการพูด เมื่อพบข้อผิดพลาด ระบบการตรวจสอบก็จะหยุดกระบวนการการผลิตภาษาทันทีและจะกลับไปยังขั้นตอนก่อนหน้านั้นในกระบวนการการผลิตภาษา ทฤษฎีที่มีอิทธิพลอื่น ๆ ก็ยอมรับแนวความคิดเกี่ยวกับการตรวจสอบด้วยเช่นกัน (MacKay, 1987; Levelt, 1989; Schade, 1999) อย่างไรก็ตาม ผลงานวิจัยชิ้นใหม่แสดงให้เห็นว่าบทบาทของกระบวนการการตรวจสอบซับซ้อนกว่านั้นมาก และสำหรับผู้เรียนที่มีทักษะทางภาษาขั้นสูงและมีการผลิตคำพูดโดยอัตโนมัติอยู่ในระดับสูงขึ้นไป ระบบการตรวจสอบจะทำการแก้ไขอย่างแอบแฝงมากขึ้น

คำสำคัญ: ระบบการตรวจสอบแก้ไข; การตรวจสอบแก้ไข; Krashen; Levelt; การรับรู้ภาษาที่สอง; การผลิตคำพูด

Introduction

In this paper I will give a survey about research on monitoring; this research comes from different fields of research. Even though Krashen, researching in the field of Second Language Acquisition, published his monitoring hypothesis in 1982, he did not yet have a comprehensive model on which he could base his hypothesis. Levelt, researching on monolingual Language Production, developed such a model in 1989. For the sake of clarity the paper breaks down chronological order and starts with Levelt's comprehensive model of speech production first to give a solid basis for discussion. Krashen's hypothesis then is applied to gain

insight into monitoring processes when acquiring a second language. More recent research then will conclude this paper.

Levelt's model of speech production

Levelt's (1989) model of speech production is up to now one of the most influential. It rests upon a firm empirical basis. This empirical data though mainly consists of speech-error data from adult L1 speakers; a direct transfer to second language contexts hence seems uncertain. The three main components of his model, namely conceptualizer, formulator, and articulator, process language in a unidirectional, incremental way. Since the model is based on data from monolingual adult speakers it is quite static, insofar that first or second language acquisition or the development of the interlanguage is not considered.

The architecture of Levelt's speech production model consists of several components that he calls "relative autonomous specialists" (Levelt, 1989, p. 22). These components work in a modular manner, since they only accept the input from a higher level. The results of each level are passed on in a top-down process in which the single components work autonomously. In other words, they do not have access to processes within other components. For example, the formulator, which is responsible for grammatical encoding, will only transfer conceptual structures into grammatical structures without regarding intonation of the whole sentence or considering the next speech intention (by this you mean the pragmatic function?), since these steps are part of the processes within the articulator or the conceptualizer respectively. The components work incrementally, so those parts of the output of one level, which are already processed, are immediately passed down to the next component without waiting for the whole element to be processed. Language production in this way is necessary in order to explain the high speed with which oral language production is processed. The

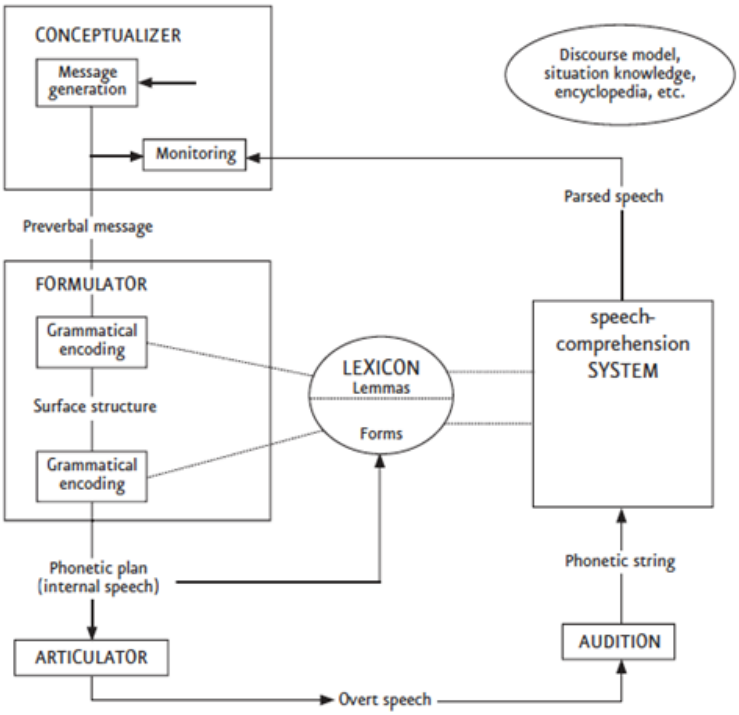
components of Levelt's model are the conceptualizer (for conceptual processing), the formulator (grammatical and morphological processing) and the articulator (phonetic processing). Levelt claims that language production is lexically driven (Levelt, 1989, p. 181). Even though the above mentioned components process structures autonomously, they do this based on information retrieved from the mental lexicon, where lexemes (word forms) and lemmas (functional values) are stored. Levelt refers here explicitly to Incremental Functional Grammar (LFG) and describes the importance of lexical entries as follows:

> *The claim that grammatical encoding is lexically driven implies that the encoding operations are largely controlled by the grammatically properties of the lemmas retrieved. It does not mean that lexical elements are procedures.* (1989, p. 236)

These procedures are carried out by above mentioned components. Levelt further distinguishes between declarative and procedural knowledge. The mental lexicon holds declarative information about word forms and their functions and supplies information to the formulator. This is complemented with other kinds of declarative knowledge, e.g. world knowledge from long time memory or knowledge about specific types of discourse; this knowledge is retrieved by the conceptualizer.

The components, on the other hand, contain only procedural knowledge. Levelt distinguishes three levels: The conceptualizer generates speech intentions, which are handed down in pre-verbal form to the formulator for grammatical encoding. The formulator then will transform the pre-verbal message into a phonetic plan. The transformation happens in two steps, first by grammatical encoding and then by phonological encoding. Finally the articulator will execute the phonetic plan and produce overt speech. The model is presented in figure 1:

Figure 1: Levelt's model of speech production



Between the individual components there is no bi-directionality; speech production is described as strict top-down processing. Nevertheless, there are certain loops incorporated. There is the possibility of feedback so that while producing language, a speaker monitors whether the utterance makes sense. This is represented as an arrow that feeds back to the Conceptualizer. This monitoring takes place through the comprehension system, which is not described in detail since Levelt focuses on speech production and not reception.

Monitoring in Levelt's model

It is worth taking a closer look at how Levelt describes monitoring. The Monitor monitors speech production and encoding processes and it stops production and initiates repairs as soon as it detects errors in the output. Levelt locates the monitor within the speech comprehension system and it monitors not only overt speech but also the phonetic plan, which is alternatively described as inner speech. It seems quite sensible to assume that inner and external speech should use the same components and resources for production, since the relation between inner and external speech had been of interest for psychologists since the works of L.S. Vygotzky in the 30´s. Levelt assumes perceptional loops that enable the monitor to detect errors – for overt speech Levelt postulates an external loop (basically a detour through the ear) and for inner speech an internal loop. An error in the phonetic plan thus is detected via the internal loop, and error in overt speech is detected via the external loop. Repair procedures are initiated as soon as an error is detected regardless of completing lexical, phonological or syntactical units. The main rule for repairs is:

> "Stop the flow of speech immediately after detecting trouble". (Levelt, 1989, p. 478)

Only correct words are excluded from this rule since their integrity is respected. Repairs are often introduced by using certain "editing expressions" (Levelt, 1989, p. 498) such as *-er /no, sorry / that is*). These expressions are usually not used when an expression is considered not appropriate. Levelt describes error repairs as conservative. The erroneous phrase is re-started within the boundaries of the intended syntactical structure, with the new phrase substituting the erroneous one (as for example in "*take the green –er the red one*"). Repairs of items that are considered not appropriate are innovative, usually triggering re-phrasing of the whole syntactical structure.

On the other hand there are repairs without stopping or re-starting, but these are causing ungrammaticality and impair the well formedness of phrases, as in "that´s the only thing he does is fight" or in "He conquered Babylon, the great Alexander" (both examples found in Levelt, 1989, p. 496). Unfortunately Levelt does not explain the procedures within the monitor in detail, but instead refers to the conceptualizer as the main component (1989, p. 14):

> "The main work is done by the conceptualizer, which can attend to internally generated messages and to the output of the speech-comprehension system (i.e., parsed internal and overt speech)".

Speakers are able to monitor all aspects of their speech production, yet speakers usually focus on selected aspects of their speech, so that certain other errors may remain undetected. It is important how much attention is directed to these aspects, and moreover there is evidence that even within phrases there are various degrees of monitoring. Levelt (1983) could show that usually the monitor seems to be more active towards the end of a phrase. Levelt assumes that at the beginning of a phrase more attention is required for planning, whereas this attention is set free for monitoring towards the end of a phrase.

Although Levelt's model might look rather complex, O' Grady et al. (1996, pp. 459-460) call it a "great simplification" of what might actually occur in the mind during language processing. This may be unjustified, since the very idea of a model is the simplification of complex procedures and the hope to find out more about its subject through falsification of the model. On the other hand Levelt excludes language learning and second language production as well as bilingual speech production, code switching, interlanguage and so on.

As speakers both in L1 and L2 evaluate their output in terms of form, content, and situation, they require monitoring.

Apart from its corrective function in ongoing speech, monitoring also plays a crucial role in language acquisition (cf. Clark, 1982). Although most researchers would concede monitoring processes in speech production, there is only little agreement upon their nature.

The Monitoring Hypothesis from Krashen

Krashen's "Monitoring Hypothesis" is well known, but has fallen out of favor with SLA researchers for obvious reasons as we will see. Nevertheless it is a fruitful attempt to explain the role of monitoring in Second Language Acquisition and addressed issues that later research would try to clarify. According to Krashen, learners of a second language develop two separate and autonomous linguistic systems. One is the acquired system, which consists of unconscious rules related to language. The other system is the learned system, consisting of explicitly learned, conscious rules about the language. Krashen puts the acquired system in a superior position regarding language knowledge and performance, and he attributes a far less important role to the explicit conscious system. He claims that this conscious system can be called upon under specific circumstances to monitor the language output. These conditions are as follows: (1) the learner has to know the rule to apply, and (2) the learner must have time to apply the conscious rule. Krashen claims that under normal speaking conditions learners do not have the time to monitor their output. Moreover, learners generally focus their attention on meaning and not on form. Thus, the monitor is generally not engaged but may show up more in writing than in speaking.

Krashen's opinion may arise from several misunderstandings. Firstly, Krashen is not precise about how to detect monitoring activities. Editing and overt repairs are easy to detect in language production, but especially the frequency and the distribution of self-repairs are suited to gain detailed insights into monitoring processes. It seems

important to also consider covert repairs, hesitation phenomena such as filled and unfilled pauses, repetitions, and drawls as well as prosodic features. These may as well indicate monitoring processes. Furthermore, it may be worthwhile to focus on the position and timing of cut-offs, false starts etc. Secondly, it is evident that a speaker is able to monitor all aspects of his speech production, e.g. situational, pragmatic and semantic aspects, not only formal correctness. Thirdly, the distinction between learning and acquisition has become rather difficult, since both concepts nowadays are rather used to signify controlled learning processes in a an educational setting on the one hand and uncontrolled input situations in a second language environment on the other hand. Despite its common popularity this concept has many shortcomings since in natural situations a strict separation between learning and acquiring is impossible. To assume two distinctively different modes of language processing in SLA is highly inefficient. In the end, it seems that Krashen's monitor is not only quite useless but also takes sociolinguistic concepts and applies them misleadingly to cognitive processes.

Still, many observations are correct and useful. There is evidence that most L1 speakers focus more on meaning than on linguistic forms (Aguado, 2003, p. 18). It seems to be more important to speak fluently and understandably than to produce grammatical flawless sentences. Bardovi-Harlig & Dörnyei (1998) claim that L1 and L2 speakers may not repair all mistakes, even though these mistakes are detected, simply because they consider these mistakes non-threatening and not crucial for the success of their communicative interaction. In the case of L2 learners nevertheless the inability to repair mistakes must be considered as a reason as well. In order to better understand the language production behavior of L2 learners it may be important to apply introspective and retrospective methods of data collection, since the learner may be able to shed light on

certain conscious aspects of her or his language production.

Monitors for inner speech and overt speech

Experimental Data so far shows that at least 2 monitors may exist – one for overt speech and one for inner speech. Hartsuiker & Kolk (2001, p. 144) refer to error repairs found in Levelt (1983) and conclude from these that it is necessary to assume a monitor for inner speech, for example in:

"take the v/ horizontal line"

Analysis of this repair shows that the speaker originally intended to say "vertical". The break-off happens much too quick to be attributed to overt speech monitoring, since the time used for uttering /v/ takes only about 70ms; hence it is necessary to assume a second monitor for inner speech. Hartsuiker could further show that in general inner monitoring is superior to overt speech monitoring. This might be explained by the assumption that overt speech monitoring is processed only through the receptive components whereas inner speech monitoring may use resources from reception and production components of the speech system. Yet the findings from Hatsuiker are not that easy to be interpreted. Inner speech monitoring mainly reports word errors, but phonetic errors are mainly detected by the overt speech monitor. Moreover detection depends on the size of the respective error.

Schade (2003) claims that especially the weaker reception based monitor plays an important role in improving second language skills, whereas the inner speech monitor rather inhibits speech production. Van Hest (1996) could show that speakers monitor less in their native language than in a second language. Only when speakers get to a higher level of second language competence, the frequency of observed monitoring will decrease.

Even though it is not always clear if the production- or the reception based monitor detects an error, there are important differences in those monitors. Whereas the

production-based monitor seems to work largely on an unconscious the reception based monitor seems to work on a more conscious level. Schade infers these from the ability of L2 speakers to reflect on their repairs, e.g. in (example in German):

> Speaker: "und&eh eh:" (Pause) "an oder auf" (Pause: lacht) "ich weiss nicht". (Schade, 2003, p. 112) [translation: "and&eh eh:" (break) "at or on" (break: laughs) "I don't know"]

Since the speaker is commenting on what has happened – her insecurity about the right preposition - it is obvious that she is aware of her monitoring. A case like this shows the work of the reception based monitor and indicates that reception based monitoring is consciously accessible. What is more important, in the continuation of the dialog the partner encourages the speaker to use the second solution ("auf") and the speaker completes the sentence in its correct form, therefore playing also a role in the acquisition of the L2.

The Monitor and Second Language Acquisition: Qualitative Findings

Another study that aims at a better understanding of the monitoring process is Kormos (2003). This study uses retrospective interviews as well as quantitative data. As Schade has shown above, monitoring processes may be consciously accessible. Kormos focused on self-repair behavior of two Hungarian learners of English and used a quantitative and qualitative analysis of their output and detailed retrospect comments they made on their performance. One (male) learner was on pre-intermediate level and had been studying English for 3 years. The other (female) learner was upper-intermediate level and had been learning English for 4 years.

Kormos used first a quantitative and then a qualitative approach. First, she found that the male learner on pre-intermediate level repairs in general happened/occurred/came about twice as frequently as the learner on upper-intermediate level; also error repairs were more frequent. The male learner seemed

to be more conscious of grammatical errors, since he corrected nearly 30 %, whereas the female learner on upper-intermediate corrected only about 18 % of her faulty grammar. This confirms Van Hest (1996) who found that pre-intermediate students make more repairs than advanced learners.

Even though this may suggest that advanced learners monitor less than pre-intermediate learners, a closer look reveals a different picture. As Kormos found out through qualitative analysis the upper-intermediate learner monitors covertly while the pre-intermediate learner makes no covert repairs at all. Here is one example of her covert repairs with an informative retrospective comment:

> Margit: Do you *er* change your date
>
> Retrospective comment: As I was saying this sentence, I noticed that it was incorrect, but I did not want to repeat the whole sentence again. I was told by my teachers that repeating a sentence sounds strange, and it matters more that you produce the sentence fluently than the fact that it contains an error. (Kormos, 2003, p. 124)

Kormos concludes that the upper-intermediate speaker has reached a higher level of automaticity in her language production and her encoding processes. Her speech is much faster than that of the male speaker. The high degree of automaticity allows her to run parallel cognitive processes and to intercept errors before they are actually audibly produced. The female speaker's attitude towards errors is remarkable as well. She explicitly states that fluency is more important to her than accuracy so she focuses more on fluency and meaning than on accuracy. The male speaker in comparison focuses on accuracy even when his utterances are perfectly understandable. Kormos points out that the distribution of attention in these two learners differs considerably. She identifies a low level of automaticity as the reason why covert repairs are missing in the pre-intermediate speaker's output, while

a high level of automaticity allowed the upper-intermediate learner to often succeed in correcting errors before they were articulated.

This qualitative analysis shows that monitoring processes rely on at least two more cognitive processes: automaticity and attention. Research in experimental psychology has shown, that a) attention can be directed, which found its expression in the "spotlight" metaphor, and b) that automaticity signifies cognitive processes that do not require attention, are fast and difficult to control. For a comprehensive survey see Engelkamp & Zimmer (2006). The interdependency of monitoring, attention and automaticity has not yet been explored and remains a desideratum for further research.

Conclusion

Although most researchers would concede monitoring processes in speech production and reception, there is only little agreement upon their nature. Krashen's (1982) monitoring hypothesis drew attention to the existence of a monitoring device and its importance for learning a second language. Error detection and subsequent self-repairs are only possible if there is a specific control device, so Levelt (1989) incorporated such a device in his language production model but claimed two different loops for inner and overt speech. Schade (2003) claims two different monitors, one production-based and another one reception-based and discusses the role of these monitors in the process of L2 acquisition. Kormos (2003) findings imply that the monitor works differently depending on the level of the second language. Based on the oral output of two learners her study suggest that with advanced language skills and a higher degree of automaticity the monitor has enough time to make covert repairs before errors are produced. Moreover, her findings put monitoring in context with other cognitive processes like attention and automaticity as well as conscious decisions of the individual learner which aspect of their

speech production to monitor.

Future research therefore should focus on the relation between monitoring, attention and automaticity. Even though automaticity has been researched in the context of fluency there is little knowledge about its role in monitoring. Generally it is assumed that monitoring needs attention as a cognitive resource. The concept of different monitors or changing ways of processing covert and open repairs demands a more detailed explanation of how attention is directed. Nevertheless it is obvious that monitoring is a pre-condition for learning or acquiring a language. The monitor indicates an error and enables the speaker to consciously notice errors in L2 production and repair errors; monitoring thus allows for language learning in spite of erroneous language output.

Krashen focused mainly on the learners input as a means of learning a second language. By contrast, most of the more recent studies assume that the learners' oral output plays a more important role in their acquisition of L2 competence. For instance, it is the basis for corrective feedback, it allows for syntactic processing, and it promotes the automaticity of speech production processes. Oral output is a precondition for the active testing of hypotheses and the acquisition of discourse competence, and especially future research into the relation between output, attention, automaticity and monitoring promises fruitful research findings.

Krashen's hypothesis was not based on a validated model of Second Language Acquisition, which may explain why it has fallen out of favor. Levelt developed such a conclusive model for monolingual speech production firmly based on empirical data. In order to fully understand the role of monitoring in First and Second Language *Production* and to understand its role in Second Language *Acquisition* it is necessary to modify this model and to unite research findings from different fields of research into a more appropriate model.

Literature

Aguado, K. (2003). Kognitive Konstituenten mündlicher Produktion in der Fremdsprache: Aufmerksamkeit, Monitoring und Automatisierung. *Fremdsprachen Lehren und Lernen, 32,* 11-26.

Bardovi-Harlig, K., & Dörnyei, Z. (1998). Do language learners recognize pragmatic violations? Pragmatic vs. grammatical awareness in instructed L2 learning. *TESOL Quarterly, 32*, 233-262.

Clark, E. V. (1982). Language change during language acquisition. In M. E. Lamb & A. L. Brown (Eds.), *Advances in Developmental Psychology* (pp. 173-197) (vol. 2). Hillsdale, NJ.

Engelkamp, J., & Zimmer, H. D. (2006). *Lehrbuch der kognitiven Psychologie*. Göttingen: Hogrefe.

Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology*, *42* (2), 113–157.

Kormos, J. (2003). Attention and monitoring in a second language: A qualitative analysis. *Fremdsprache Lehren und Lernen*, *32*, 116–128.

Krashen, S. D. (1982). Second language acquisition and second language learning. New York: Prentice-Hall International.

Levelt, W. (1983). Monitoring and self-repair in L2. *Language Learning*, *49*, 303-342.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer.

O' Grady, W., Dobrovolsky, M., & Katamba, F. (1996). *Contemporary linguistics.* London: Longman.

Schade, U. (1999). *Konnektionistische Sprachproduktion*. Wiesbaden: DUV.

Schade, U. (2003). Monitoring-prozesse und ihre Rolle für die fremdsprachliche Produktion. *Fremdsprachen Lehren und Lernen*, *32*, 99–115.

Van Hest, E. (1996). *Self-Repair in L1 and L2 production*. Tilburg: Tilburg University Press.