Words, pauses, and gestures: New directions in language production research

Schiller, N.O.; Ferreira, V.S.; Alario, F.X.

LANGUAGE AND COGNITIVE PROCESSES
2007, 22 (8), 1145–1150

Psychology Press
Taylor & Francis Group

Words, pauses, and gestures: New directions in language production research

Niels O. Schiller
Leiden Institute for Brain and Cognition (LIBC), Leiden, the Netherlands, and Universiteit Maastricht, Maastricht, the Netherlands

Victor S. Ferreira
University of California, San Diego, CA, USA

F.-Xavier Alario
CNRS and Université de Provence, Marseille, France

We present a summary of recent research in the area of language production based on contributions presented at the *Second International Workshop on Language Production* (Maastricht, the Netherlands, 28–30 August, 2005). The articles included in the present special issue report on the production of words, pauses, and gestures, as well as the effects of ageing on lexicalisation processes. These papers raise a range of relevant issues for the study of language, including linguistic vs. cognitive influences on language production (F. Ferreira), the use of eye movements to assess the effect of ageing on specific aspects of cognitive performance (Belke and Meyer), how gesture can be used to better understand cross-linguistic differences (Kita et al.), and advanced theory development in the well-studied domain of single-word production (Roelofs; La Heij et al.; and Roelofs).

Correspondence should be addressed to Niels O. Schiller, Cognitive Psychology Unit, Department of Psychology, Faculty of Social Sciences, Leiden University, P. O. Box 9555, NL-2300 RB Leiden, the Netherlands. E-mail: nschiller@fsw.leidenuniv.nl

The *Second International Workshop on Language Production* held in Maastricht (The Netherlands) in August 2005 was made possible by financial support from the Faculty of Psychology of Maastricht University. The scientific committee consisted, besides the authors, of Albert Costa and Martin Pickering. Niels O. Schiller was supported by grant no. 453-02-006 from the Netherlands Organisation for Scientific Research (NWO).

http://www.psypress.com/lcp DOI: 10.1080/01690960701491415

WORDS, PAUSES, AND GESTURES: NEW DIRECTIONS IN LANGUAGE PRODUCTION RESEARCH

This special issue on language production contains selected articles from contributions presented at the *Second International Workshop on Language Production*, organised by the Psycholinguistic Research Group of the Department of Cognitive Neuroscience at Maastricht University in Maastricht (the Netherlands) in August 2005. This workshop was the second annual meeting, following the successful workshop in 2004 in Marseille (see special issue in *Language and Cognitive Processes*, volume 21, issues 7/8). The meeting in Maastricht brought together researchers with different perspectives on language production who approach the field using different methodologies. Accordingly, this special issue comprises articles that address original questions within the field of language production (such as age effects on language production performance, the origin of the gestures that accompany speech, or the relative contributions of competence and performance to speaking), as well as articles that address more central issues (e.g., the computational modelling of lexical access). This variety of topics and methods reflects the maturity and diversity of the field of language production (for a general discussion of the state of the art of language production we refer the reader to Alario, Costa, Ferreira, & Pickering, 2006).

When we speak we produce words, pauses, and gestures. The rate and speed of their production depends on a number of factors. Also, we have the intuition that, as we age, we produce patterns of words, pauses, and gestures in quantitatively different ways compared with younger speakers. The articles in this special issue each address these topics.

This special issue starts with an article by Fernanda Ferreira (2007) in which she investigates a classic question in psycholinguistics: How much of a speaker's behaviour can be attributed to respecting linguistic rules and how much is due to cognitive constraints of on-line processing? This is an instantiation of the classic distinction between competence and performance. F. Ferreira considers the example of the pauses a speaker produces within a sentence. She makes a proposal for distinguishing prosodic effects, arising from the linguistic requirements that apply to the utterance the speaker is producing, from planning effects that arise from the need to plan ahead and retrieve from memory a certain amount of upcoming linguistic material. In her article, F. Ferreira provides a critical review of models developed to account for the likelihood with which pauses arise in the course of sentence production (e.g., Gee & Grosjean, 1983; Watson & Gibson, 2004). She then describes the results of her own work in which prosodic constraints and planning performance are shown to differentially affect pause duration. Prosodic constraints are thought to come from linguistic material prior to

the pause, whereas planning performance constraints depend on material following the pause boundary. The article emphasises the importance of making use of this distinction when modelling prosodic production, and provides some suggestions on how to achieve this.

The second article of this special issue is by Eva Belke and Antje Meyer. These authors registered eye movements to investigate lexicalisation processes in younger (mean age: 20 years) and older speakers (mean age: 68 years). Although eye movements may be less intuitively linked to speech planning processes, they have been shown to be highly correlated with speech output. For instance, speakers describing visual displays usually gaze at objects until they have assembled the phonological form of their names (Griffin, 2001; Meyer, Sleiderink, & Levelt, 1998). When an object name is less familiar, longer, or for some other reason more difficult to retrieve and/or encode, speakers look longer at the object than when the object name corresponds to a high-frequency, short name that is easy to retrieve (Griffin & Bock, 2000; Meyer & Van der Meulen, 2000; see overview in Griffin, 2004). In their study, Belke and Meyer (2007) found that older participants exhibited significant slowing in multiple object naming, reflected by slower speech and longer gazes, whereas in single object naming, there were no differences in performance between the two age groups. More generally, older speakers do not seem to exhibit a substantial decline in their ability to name common objects. Rather, since they are slower than younger speakers when several objects must be named in a row, Belke and Meyer conclude that older speakers may allocate more processing capacities to speech monitoring processes (see Postma, 2000 for an overview) than younger speakers, which may slow down speech planning processes. By contrast, no age-related slowing of lexical retrieval processes seems apparent (Burke & Shafto, 2004).

This is followed by an article by Sotaro Kita and colleagues investigating spontaneous gestures accompanying speech production. An interesting but not yet heavily investigated phenomenon in language production research is gesture: The movement or positioning of parts of the body that can convey meaning and that accompanies or replaces verbal communication. Gesture is interesting for at least two reasons: First, it is an important phenomenon to understand in its own right, both in terms of its communicative consequences and its cognitive basis. Second, for users of spoken language at least, gesture involves tightly coordinated interactions between systems in different modalities. Given current trends in psychological research, both of these issues are ripe for vigorous investigation. A trend gaining momentum in language research is to investigate language in more natural settings, and to investigate aspects of language use that are observed almost exclusively in natural discourse (e.g., disfluency). Gesture is just such a feature.

Kita, Özyürek, Allen, Brown, Furman, and Ishizuka (2007) investigate the relationship between gesture and spoken production by looking at a

previously discovered relationship between gesture and speech. Specifically, Özyürek and Kita (1999) and Kita and Özyürek (2003) showed that in languages such as Japanese or Turkish that separate the description of the path along which an object moves and the manner with which that object moves (so that they effectively say 'the ball went down the hill rolling'), speakers also separate the accompanying gestures for those elements, such that speakers will first gesture down the path, and then will indicate the rolling manner. However, in languages like English where these elements are conflated (e.g., 'the ball rolled down the hill'), so too are the gestures. Kita et al. look to determine whether the cognitive basis of this pattern is a static, schema representation acquired through general language experience, or whether it is instead due to a dynamic, on-line interaction between the processes that organise speech and those that organise gesture. They discovered evidence for the latter: When speakers of English produce utterances that conflate path and manner, their gestures too are conflated, but when they produce path and manner separately, their gestures are also separated. This not only helps us to better understand gesture, but to better understand the cognitive basis of speech-body interactions more generally.

Finally, this special issue is completed by a theoretical discussion between Ardi Roelofs and Wido La Heij and his colleagues on what has arguably been the most investigated topic in language production research over the last 20 years or so, namely lexical selection during single word production (Goldrick, 2007). The discussion focuses on the modelling of performance in picture-word-interference paradigms, where participants are asked to produce a word (e.g., in response to a picture or as a translation) while they ignore a distractor word whose properties are manipulated. Roelofs (2007a) provides a critical and sharp analysis of the name retrieval model proposed by La Heij and colleagues (Bloem & La Heij, 2003; Bloem, Van den Boogaard, & La Heij, 2004; Starreveld & La Heij, 1996). On the basis of his detailed analysis, he criticises the model for being too simple and further argues that if the model's complexity were to be increased to meet some empirical requirements, it may lose its distinctiveness from other proposals (e.g., WEAVER++; Levelt, Roelofs, & Meyer, 1999). La Heij, Starreveld, and Kuipers' (2007) reply provides detailed counter-arguments on these points, both on empirical and theoretical grounds. La Heij et al. argue that simplicity is not the sole motivation for their model, and they defend their approach which seeks to test modelling principles that account for some empirical phenomenon, rather than providing a comprehensive model of lexical access in its entirety. They point to core differences between their proposal and alternative models. Finally, Roelofs' (2007b) rejoinder questions some aspects of their reply.

It is important to highlight the scope of this discussion because it is detailed and makes extensive use of evidence from the specific picture-word-

interference paradigm (though other types of evidence such as speech errors are drawn on as well). A commonly heard argument against this kind of detailed discussion is that the goal of psycholinguistic research should not be to provide a model of the picture-word-interference task (or the lexical decision task, or any task). The goal should rather be to understand the mechanisms underlying language ability in any possible context. This argument is important, yet it should not be confused with another argument stating that, because a task has been investigated for over two decades, the available evidence and theoretical discussion have become too complex and the task should be dismissed. In attempting to circumvent this complexity by using a novel task, which is not substantially simpler, research efforts are bound to develop a similar task-based knowledge over the years, ultimately suffering from the same criticism. In short, then, the detailed modelling efforts made to account for the results observed in the picture-word-interference task, as well as other complementary lines of evidence, reflect how much progress has been made in understanding the mechanisms involved in lexical selection. Unless proven otherwise, the complex pattern of data from this popular task provides unavoidable constraints on models of language production.

Our view is that the articles presented in this special issue move the field of language production forward. These articles deal with detailed computational implementations, the interaction of linguistic and cognitive constraints and with the rather novel topic of ageing and lexical processing which will surely be among central issues in the next years for research on language production.

REFERENCES

Alario, F.-X., Costa, A., Ferreira, V. S., & Pickering, M. J. (2006). Architectures, representations and processes of language production. *Language and Cognitive Processes*, *21*, 777–789.

Belke, E., & Meyer, A. S. (2007). Single and multiple object naming in healthy ageing. *Language and Cognitive Processes*, *22*, 1178–1211.

Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language*, *48*, 468–488.

Bloem, I., Van den Boogaard, S., & La Heij, W. (2004). Semantic facilitation and semantic interference in language production: Further evidence for the conceptual selection model of lexical access. *Journal of Memory and Language*, *51*, 307–323.

Burke, D. M, & Shafto, M. A. (2004). Aging and language production. *Current Directions in Psychological Science*, *13*, 21–24.

Ferreira, F. (2007). Prosody and performance in language production. *Language and Cognitive Processes*, *22*, 1151–1177.

Gee, J. P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, *15*, 411–458.

Goldrick, M. (2007). Connectionist approaches to language production. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 515–530). Oxford: Oxford University Press.

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition*, *82*, B1–B14.

Griffin, Z. M. (2004). Why look? Reasons for eye movements related to language production. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 213–247). New York: Psychology Press.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*, 274–279.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*, 16–32.

Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes*, *22*, 1212–1236.

La Heij, W., Starreveld, P. A., & Kuipers, J.-R. (2007). Structural complexity is not the (big) issue: A reply to Roelofs (2007). *Language and Cognitive Processes*, *22*, 1261–1280.

Levelt, W. J. M, Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.

Meyer, A. S., Sleiderink, A., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, *66*, B25–B33.

Meyer, A. S., & Van der Meulen, F. F. (2000). Phonological priming effects on speech onset latencies and viewing times in object naming. *Psychonomic Bulletin and Review*, *7*, 314–319.

Özyürek, A., & Kita, S. (1999). Expressing manner and path in English and Turkish: Differences in speech, gesture, and conceptualization. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the Twenty-First Annual Conference of the Cognitive Science Society* (pp. 507–512). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, *77*, 97–132.

Roelofs, A. (2007a). A critique of simple name-retrieval models of spoken word planning. *Language and Cognitive Processes*, *22*, 1237–1260.

Roelofs, A. (2007b). On the modelling of spoken word planning: Rejoinder to La Heij, Starreveld, and Kuipers (2007). *Language and Cognitive Processes*, *22*, 1281–1286.

Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 896–918.

Watson, D., & Gibson, E. (2004). The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes*, *19*, 713–755.