PROCESSING LANGUAGE PERFECTLY BUT NOT AUTOMATICALLY

Janet Nicol
University of Arizona

Teresa R. Bell
University of Oklahoma

In this paper, we report the results of a preliminary study in which we explore the timing of productions of short noun phrases vs. word lists by native German speakers and advanced learners of German. The phrases/word lists consisted of a noun preceded by two adjectives that appeared in either a more usual sequence or unusual (but grammatical) sequence. Participants were asked to either produce the words as a list or as a phrase (which required the adjectives to agree with the noun in number and gender). Although both groups produced error-free utterances and showed similar response times to utter the lists, the native German speakers, but not the second language speakers, were significantly slowed by the unusual adjective order, but only when they were producing phrases.

INTRODUCTION

As second language learners quickly realize—and monolinguals rarely do—language production is a complicated business. Of course, it doesn't appear to be to monolinguals, because they have had a lifetime of practice. People spend many hours every day engaged in the task of producing sentences: generating ideas to be communicated, finding words that best fit those ideas, organizing the words into grammatically-permissible strings, specifying agreement and other grammatical features, and uttering a long and complex sequence of sounds. All this within a short time span—sentences are uttered at a rate of about 150 msec. per syllable—to meet the constraints of conversation.

When people learn a second language, they must find new words to match their thoughts, sequence words into a different order, attend to new grammatical features and generate the appropriate morphemes, and utter strings of new sounds in new combinations. When people speak a second language well enough that they produce fluent, error-free utterances, it is reasonable to suppose that they have mastered these different facets of production so that they are producing language in the same manner as native speakers. In other words, if the output of a native and nonnative speaker is identical, the underlying mechanisms are likely to be operating in identical fashion.

In this paper, we report the results of a preliminary study that uses

reaction time data to explore subtle differences between native speakers and proficient second language speakers. The findings show that second language speakers and native speakers show a markedly different pattern of performance, suggesting that the machinery that generates language in the two groups is also quite different.

Our study was modeled on research by Pechmann (1989, 1994) and Pechmann and Zerbst (1993, 1995), who examined the production of noun phrases in German. Specifically, they examined how quickly speakers were able to utter a phrase consisting of a size adjective (e.g. *little*), a color adjective (e.g. *red*), and a noun (e.g. *book*). In both German and English, there is a preferred order of adjectives: size adjectives appear before color adjectives (Martin, 1969). The reverse order is possible under particular discourse conditions (e.g. "not the blue little book, the red little book"), though this construction is obviously marked. Pechmann and Zerbst presented their German participants with lists of three words such as *rot, klein, Buch* (*red, little, book*) or *klein, rot, Buch* (*little, red, book*) and asked them to produce either a phrase, in which the adjectives would agree in gender and number with the noun, or to simply say the words as a list. They found that participants were faster to initiate the utterance when the adjectives were in the preferred order, but only when they were asked to produce a phrase. When they uttered the words as a list, there was no response time difference. This suggests that preferred adjective order facilitates language production only when speakers are engaged in producing phrases and sentences.

The purpose of our experiment was to explore whether advanced learners of German show a similar pattern. If the development of proficiency in a particular language involves the automatization of a new set of routines, then the second language learners should resemble the native speakers with respect to response time differences.

METHOD

Participants

Sixteen subjects volunteered to participate in the experiment. Eight subjects were non-native speakers (NNSs) of German, and eight were native speakers (NSs) of German. Of the eight NNSs, seven were American graduate students in the M.A. program in German Studies at the University of Arizona, and five of those seven were graduate teaching assistants and teach one to two courses of German each semester. The other NNS was a senior in the BA program in German Studies. All eight had lived in Germany or a German-speaking country for some period of time, all spoke German with advanced fluency, and all continued to use German regularly. Of the eight NSs of German, six were graduate students in the MA program in German Studies, one was in the MS program in Optical Science, and one was visiting the United States from Germany.

Materials and Procedure

Two lists of phrases were constructed that each contained six preferred orders (size adjective, color adjective, noun) and six dispreferred orders (color adjective, size adjective, noun). The twelve phrases within each list were randomized so that participants could not predict what type of phrase would appear next. Each participant was presented with both lists. For one list, participants were asked to create phrases (the *phrase* condition; e.g. *kleines rotes Buch*), and for the other list, they were asked to say the words as if they were in a list (the *list* condition). Half the participants in each group were given the list condition first, and half were given the phrase condition first. Further, half the participants in each group were given List A first and half were given List B first. Each list was preceded by five practice items to familiarize the participants with the task.

Participants were tested individually in sound resistant test booths using DMASTR software (developed by K.I. Forster and J.C. Forster at the University of Arizona). They were seated in front of a computer monitor and keyboard. When an item (a string of two adjectives and a noun, separated by spaces) appeared on the monitor, participants were asked to read the words silently and then press the right SHIFT key on the keyboard when they were ready to begin uttering the phrase. The press of the SHIFT key stopped a clock internal to the computer that had been started by the appearance of the three words. Response times were recorded into data files generated by the software. The utterances were recorded onto audiotape (and were later assessed for accuracy). Participants advanced from one item to the next with the press of a keyboard key, and so could pace themselves through the experiment.

RESULTS

First, tape recordings were assessed: participants produced no errors; hence all the reaction time data were subjected to analysis. Table One shows the mean reaction times (in milliseconds) in the list vs. phrase conditions for the two groups of speakers.

It is critical to point out here that we are interested in two comparisons: (1) the difference between list RTs and phrase RTs and (2) the difference between the phrase and list conditions for each adjective order. Note that we cannot meaningfully compare preferred vs. dispreferred orders in just the list condition or just the phrase condition for one of the groups because different sets of materials was used for each of the orders. Different stimulus materials may well differ with respect to articulatory difficulty and this would affect RTs, but would be unrelated to the questions of interest. But we can meaningfully compare phrase vs. list differences (collapsing across order type) and we can compare the difference scores.

For the native German speakers, there was no main effect of list vs. phrase. Although it took longer to initiate the production of phrases, this difference was not significant ($p > .49$). Now consider the difference scores for the Preferred vs. Dispreferred orders: there was a much RT larger difference

for the Dispreferred order in lists vs. phrases (153 msec.) than for the Preferred order in lists vs. phrases (7 msec). The difference between 153 msec. and 7 msec. is highly significant $F(1,6) = 27.27$; $p = 0.00197$. Hence, for German speakers, order of adjectives differentially affects the production of lists vs. phrases. This replicates the results reported by Pechmann and Zerbst (1995).

Table 1: Mean Reaction Times (msec.) for the Different Treatment Conditions for the Two Groups of Participants in Experiment 1

		List	Phrase	**Difference**
Native German Speakers	**Preferred Adjective Order**	1340.9	1333.5	**7.4**
	Dispreferred Adjective Order	1197.8	1351.0	**-153.2**
Non-native German Speakers	**Preferred Adjective Order**	1138.00	1882.88	**-744.8**
	Dispreferred Adjective Order	1172.50	1896.00	**-723.5**

Now let us consider the proficient nonnative German speakers. They were notably slower to initiate the production of phrases (close to 1900 msec.) compared to lists (about 1150 msec., comparable to the native speakers). This main effect is significant; $F(1,6) = 10.9$; $p = .0164$. However, in contrast to the native German speakers, there was not a hint of an interaction; adjective order had no effect whatever on the production of phrases vs. lists ($p > .88$).

DISCUSSION AND CONCLUSION

There are two major findings. First is the finding that native German speakers show an interaction between adjective order and phrase type (phrase vs. list), and second language learners do not. Second is the finding that English show a large difference between list and phrase conditions.

The fact that native speakers and nonnative speakers show different patterns of performance indicates that even when the nonnative speakers produce error-free utterances, there are underlying differences in how they prepare and execute their productions. A characterization of that difference requires some speculation about why the native German speakers show the pattern they do. In the phrase task, participants must take into account the third element (the noun) in determining the forms of the adjectives. In doing so, they may consider how strange the pair of adjectives sounds when they appear in the dispreferred order and this slows them down. In the list task, however,

they may launch directly into uttering the adjectives without first considering the specifications of the noun. To the extent that German speakers do consider how the words sound together, they may be more struck by the fact that the adjectives are uninflected than by the oddness of the dispreferred order. When the second language learners produce phrases, they may be so intent on getting the adjective inflections right, and so unpracticed at producing complex noun phrases, that they do not notice the oddness of the dispreferred order (despite the fact that the adjective preference is identical in English). In other words, they may be so focused on retrieving the correct forms—a task which clearly takes them considerably longer than it does the native speakers—that they have no time to assess how the words sound together. In the list task, like the German speakers, they may initiate the utterance prior to having interpreted the words as a group, so again, the order of the adjectives does not affect response times.

The performance difference between the two groups points to two areas of the language production system that are less than fully automatic. One is the inflection of adjectives: nonnatives appear to take considerable time to retrieve the right forms. The other is the assessment of how good the adjectives sound together in the two orders. Current models of language production (e.g. Garrett, 1984, Bock & Levelt, 1994) assume that speakers monitor their output so that ill-formed or inappropriate words and sentences do not slip out. This may be the point at which German speakers notice (even unconsciously) the peculiarity of the dispreferred order. Nonnative speakers may not have the luxury of monitoring all aspects of their output, especially if the process of inflecting adjectives to agree with the following noun demands too much time and/or resources.

This experiment provides a complement to a recent study by Guillelmon and Grosjean (2001). In their experiments with native vs. nonnative speakers of French, they presented spoken phrases containing a determiner, adjective and noun. The adjective was unspecified for gender (in its spoken form), but the determiner was either masculine, feminine or neuter. Further the determine and noun were either gender-congruent (*le joli camp*; "the<masculine> pretty camp<masculine>"), gender-incongruent (*la jolie camp*; "the<feminine> pretty camp<masculine>") or gender-neutral (*leur joli camp*; "their pretty camp"). Participants were asked to repeat the last word, and their response times were recorded. The investigators found that native speakers and "early bilinguals" (who had acquired both languages from a young age) both showed congruence effects (faster times for congruent vs. incongruent cases). Proficient late bilinguals, however, showed no congruence effect. So, despite the fact that this latter group regularly used French, knew the grammatical gender of all the nouns in the experiment and produced them with the correct determiner, they were simply not processing the cue to gender provided by the determiner in the same way as the other two groups.

Obviously, some of the processing routines required for the production and comprehension of a second language can become automatic. For example, there is evidence that a characteristic type of brainwave elicited

by semantic incongruence is triggered in nonnative speakers as well as native speakers (Ardal et al., 1990). But it appears that the processes associated with grammatical elements may not become automatic. Clearly, this is an area for further research.

REFERENCES

Ardal, S., Donald, M.W., Meuter, R., Muldrew, S., & Luce, M. (1990). Brain responses to semantic incongruity in bilinguals. *Brain and Language, 39*, 187-205.

Bock, J.K., & Levelt, W.J.M. (1994). Language production: Grammatical encoding. M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*. (pp. 945-984). San Diego: Academic Press.

Garrett, M. F. (1984). The organization of processing structure for language production: Applications to aphasic speech. In D. Caplan, A.R. Lecours, & A. Smith (Eds.), *Biological perspectives on language*. (pp. 172-193). Cambridge, MA: MIT Press.

Martin, J. E. (1969). Semantic determinants of preferred adjective order. *Journal of Verbal Learning and Verbal Behavior, 8,* 697-704.

Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics, 27,* 89-110.

Pechmann, T. (1994). *Sprachproduktion*. Opladen: Westdeutscher Verlag.

Pechmann, T., & Zerbst, D. (1993). Zur allmählichen Verfernung komplexer Nominalphrasen beim Reden. In S. Felix, C. Habel, & G. Rickheit (Eds.), *Kognitive Linguistik Repräsentationen und Prozesse*. Opladen: Westdeutscher Verlag.

Pechmann, T., & Zerbst, D. (1995). Zum Einfluß der syntaktischen Struktur auf den Zeitlauf von Sprachproduktionsprozessen. *Sprache & Kognition, 14*, 64-74.