

Syllable structure effects turn out to be word length effects: Comment on Santiago et al. (2000)

Ardi Roelofs

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Santiago, MacKay, Palma, and Rho (2000) report two picture naming experiments examining the role of syllable onset complexity and number of syllables in spoken word production. Experiment 1 showed that naming latencies are longer for words with two syllables (e.g., *demon*) than one syllable (e.g., *duck*), and longer for words beginning with a consonant cluster (e.g., *drill*) than a single consonant (e.g., *duck*). Experiment 2 replicated these findings and showed that the complexity of the syllable nucleus and coda has no effect. These results are taken to support MacKay's (1987) Node Structure theory and to refute models such as WEAVER++ (Roelofs, 1997a) that predict effects of word length but not of onset complexity and number of syllables per se. In this comment, I show that a re-analysis of the data of Santiago et al. that takes word length into account leads to the opposite conclusion. The observed effects of onset complexity and number of syllables appear to be length effects, supporting WEAVER++ and contradicting the Node Structure theory.

Theories of spoken word production differ in their claims about the role of syllables and their internal structure. Santiago, MacKay, Palma, and Rho (2000) contrasted the predictions of two types of theory. According to MacKay's (1987) Node Structure theory, henceforth NST, syllables and their internal structure are planned via a tree traversal process that operates in a top-down and left-to-right fashion. In particular, the initial syllable of a word is planned before its non-initial syllables, and syllable

Requests for reprints should be addressed to Ardi Roelofs, Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH, Nijmegen, The Netherlands. Email: ardi@mpi.nl.

I am grateful to Pim Levelt and the members of the Simple Utterance Generation group of the Max Planck Institute for Psycholinguistics for helpful discussions.

onsets are planned before syllable rimes. Within syllable onsets, initial speech segments are planned before non-initial segments, and within rimes, nuclei are planned before codas. At each node in the tree, a binary decision is taken (e.g., to plan *de* rather than *mon* first in producing the word *demon*). The number of decisions that lead to the left-most, bottom-most node (i.e., the first segment) determines the production onset latency. This predicts that the production latency should be longer for disyllabic (e.g., *demon*, *dragon*) than for monosyllabic words (e.g., *duck*, *drill*), and longer for words with a complex onset (e.g., *drill*, *dragon*) than with a simple onset (e.g., *duck demon*). The examples *duck*, *demon*, *drill*, and *dragon* are taken from the paper of Santiago et al. (2000). Because traversing the phonological tree is a sequential process, effects of onset complexity and number of syllables should be additive. And because the nucleus and coda (e.g., the /ɪ/ and /l/ of *drill*, respectively) are not involved in the decisions needed to reach the first segment, the complexity of the nucleus and coda should have no effect.

Other models make different predictions. In particular, Santiago et al. compare the predictions of NST with those from the theory proposed by Levelt and colleagues (Levelt, 1992; Levelt, Roelofs, & Meyer, 1999), which is computationally implemented in the WEAVER++ model (e.g., Roelofs, 1996, 1997a,b). In WEAVER++, the speech segments of a word are activated in parallel and they are syllabified in a sequential left-to-right manner. After syllabification, phonetic syllable programs are accessed from memory and stored in an output buffer as a linear string. Only when the buffer contains the phonetic programs for one phonological word can articulation begin. This model predicts latency effects of word length in terms of number of speech segments, but it predicts no effects of onset complexity and number of syllables per se. Furthermore, the model also predicts a lack of an effect of nucleus and coda complexity as such, although having more coda segments should increase latency.

Santiago et al. ran two carefully designed experiments to test these predictions. Experiment 1 tested for effects of onset complexity and number of syllables. Experiment 2 tested the same predictions with identical materials but different participants and, in addition, tested for effects of nucleus and coda complexity. The experiments controlled for articulatory factors by matching the first segment of words pairwise between conditions (e.g., *duck* in the simple onset condition was paired with *drill* in the complex onset condition) and word frequency was also matched. The materials in each experiment were tested in three different tasks: delayed naming, picture naming, and name-picture matching. On each trial in the delayed naming task, participants were first given a printed word. A second later, the corresponding picture was presented, which had to be named after a “go” signal. This task allowed participants to practice

the production of the picture names. In addition, it provided a check on the effectiveness of the control for articulatory factors. The picture naming task was the crucial task, and simply measured the production latencies for naming the pictured objects. Naming latencies do not only reflect properties of the planned words but also the influence of perceptual and conceptual factors (e.g., some pictured objects are more difficult to recognise than others). To control for such perceptual/conceptual factors, the experimental materials together with fillers were also tested in a name-picture matching task. On each trial, a word was presented a second before a picture, and upon picture presentation participants had to indicate by pressing a *yes* or *no* button whether the word was the name of the picture. Unknown to the participants, the critical, experimental items all required *yes* responses. The central measure of interest in both experiments was the mean picture naming latency in each condition corrected for the influence of perceptual/conceptual factors. This measure was obtained by subtracting the name-picture matching latency from the picture naming latency, henceforth referred to as the *corrected* latency following Santiago et al.

Experiment 1 showed that the corrected naming latencies were longer for words with two syllables (e.g., *demon*, *dragon*) than one syllable (e.g., *duck*, *drill*), and longer for words beginning with a consonant cluster (e.g., *drill*, *dragon*) than a single consonant (e.g., *duck*, *demon*). Experiment 2 replicated these findings (with the caveat that number of syllables was now not tested in the name-picture matching task) and, in addition, showed that the complexity of the syllable nucleus and coda had no effect.

According to Santiago et al., these findings support the NST and refute models such as WEAVER++ that predict effects of word length but not of onset complexity and number of syllables per se. Santiago and MacKay (1999) reject WEAVER++ on the basis of the same data (then still unpublished) in a peer commentary on a BBS target article that explained WEAVER++ (Levelt et al., 1999). They argue against an interpretation of their data in terms of word length, but curiously, their experiments did not explicitly control for this factor. An inspection of their materials reveals, however, that onset complexity and number of syllables systematically vary with the length of the words in terms of number of speech segments. Like in the examples *duck*, *demon*, *drill*, and *dragon*, the disyllabic words in the experiments (e.g., *demon*, *dragon*) contained more segments than the monosyllabic words (e.g., *duck*, *drill*). And the words with the complex onsets (e.g., *drill*, *dragon*) contained more segments than the words with the simple onsets (e.g., *duck*, *demon*). The Appendices A and C of Santiago et al.'s (2000) paper list the materials and the mean corrected latencies for the relevant words in their experiments. This allowed me to re-analyse their data while controlling for word length. In particular, I ran analyses of variance with word length as covariate.

A RE-ANALYSIS OF THE DATA OF SANTIAGO ET AL. (2000)

Experiment 1

Appendix A of Santiago et al. (2000) shows that for the monosyllables, the words are longer in the complex onset condition than in the simple onset condition for 9 of the 14 matched word pairs. For the disyllables, this word length difference between the simple and complex onset condition even holds for 12 of the 14 matched word pairs. In these counts, I have treated diphthongs as single segments, as suggested by the outcomes of Experiment 2 of Santiago et al., where no effect of syllable nucleus complexity was obtained. Counting diphthongs as two segments gives almost identical results, namely 9 out of 14 for the monosyllables and 11 out of 14 for the disyllables. The word length difference between the simple and complex onset condition is statistically highly significant, $F(1, 52) = 42.87$, $MSE = 0.35$, $p = .000$. The Appendix A of Santiago et al. also shows that the disyllabic words in the experiments contained more segments than the monosyllabic words. Again, the difference in word length between conditions is highly significant, $F(1, 52) = 215.39$, $MSE = .35$, $p = .000$. In their statistical analyses of the production latencies, Santiago et al. obtained an effect of number of syllables, which I replicated with $F(1, 52) = 6.39$, $MSE = 10068$, $p = .015$, and a “large but unreliable onset complexity effect” (as Santiago et al. call it, p. 22), with $F(1, 52) = 1.81$, $MSE = 10068$, $p = .184$ (all statistics reported in this comment are from my own re-analysis of the data).¹ And together, the effects were additive. However, running the ANOVA with word length as a covariate (e.g., Winer, Brown, & Michels, 1991) shows that any suggestion of effects completely disappears when length is controlled. For the number of syllables, $F(1, 51) = 0.76$, $MSE = 10252$, $p = .387$, and for the onset complexity, $F(1, 51) = 0.66$, $MSE = 10252$, $p = .420$. This shows that the effects of onset complexity and number of syllables can simply be attributed to differences in word length. The syllable structure effects turn out to be word length effects.

It appears that the relation between length and latencies in the experiment is almost linear. The mean latencies for the words with 2, 3, 4, 5, 6, 7, and 8 segments are 226, 251, 283, 329, 313, 372, and 382 ms, respectively. Figure 1 plots the mean corrected latency against word

¹ My F values are identical to those of Santiago et al. However, Santiago et al. (2000) report mean latencies by participants in their tables and figures, whereas I report them, by necessity, by items. This accounts for the minor discrepancies between their and my descriptive statistics.

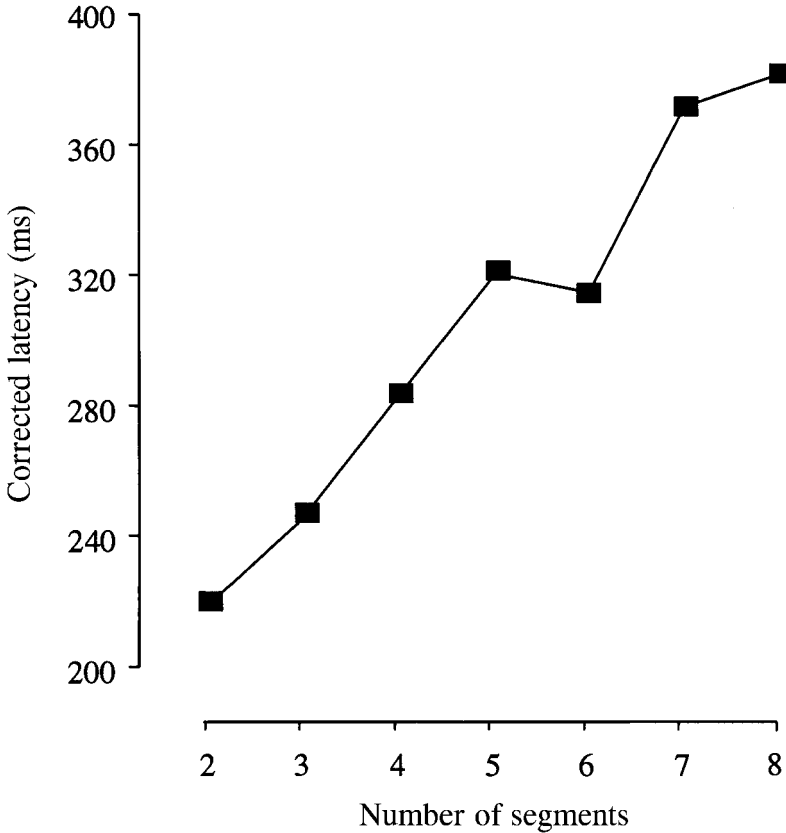


Figure 1. Corrected production latency (in ms) as a function of word length in terms of number of speech segments as measured by Santiago et al. (2000) in Experiment 1.

length. The Pearson product-moment correlation between length and mean latency is $r = .97$, with $p = .000$ ($N = 7$), and for the individual words it is $r = .34$, with $p = .009$ ($N = 56$). The correlation between length and latencies is not perfect. However, it is difficult to maintain that latency does not increase with word length in terms of number of segments, as Santiago et al. (2000) claim to have observed and NST predicts.

Note that word length also explains why the effects of onset complexity and number of syllables are additive. The disyllabic words are longer than the monosyllabic words, and within each of these two sets, the words with the complex onsets are longer than those with the simple onsets. Hence, one obtains main effects of onset complexity and number of syllables, and together, the effects are additive.

Experiment 2

Appendix C of Santiago et al. (2000) shows that for 17 of the 18 matched word pairs it holds that the words are longer in the complex onset condition than in the simple onset condition. This difference is highly significant, $F(1, 30) = 111.11$, $MSE = 0.10$, $p = .000$. In their analyses of the production latencies, Santiago et al. obtained an effect of onset complexity, $F(1, 30) = 4.19$, $MSE = 12205$, $p = .049$. (To increase power, I ran the analysis on the complete set of words, whereas Santiago et al. performed two analyses with less power each, contrasting onset with nucleus complexity and onset with coda complexity, respectively.) However, running the ANOVA with word length as covariate shows that the onset complexity effect completely disappears when word length is controlled for, $F(1, 29) = 1.24$, $MSE = 12608$, $p = .275$. This shows that the effect of onset complexity in the second experiment can also simply be described as an effect of word length.

The relation between length and latency is, again, almost linear. The production latencies for the words with 2, 3, 4, and 5 segments are 240, 233, 301, and 347 ms. The correlation between length and mean latency is $r = .93$, with $p = .070$ ($N = 4$), and for the individual words it is $r = .35$, with $p = .036$ ($N = 36$). As can be seen, there is one clear exception to a linear relation, namely words of 2 and 3 segments take, on average, about equally long to produce. It appears that this deviant finding can be attributed to a single item, namely *tea*, which is included in the CVC condition by Santiago et al. This word had a corrected latency of 404 ms. All other words in the CVC condition had latencies in the range of [107, 260] ms. There are two other 2-segment words in the experiment, listed under CVV(C), which had latencies of 125 and 191 ms. Thus, the latency of *tea* is clearly exceptionally high. When one excludes the outlier *tea* from the analysis, the correlation between length and mean latency becomes $r = .99$, $p = .005$ ($N = 4$), and for the individual words it becomes $r = .45$, $p = .007$ ($N = 35$). Figure 2 plots the mean corrected latency against word length for the remaining 35 items. As can be seen, latency increases with word length, and the relation is almost perfectly linear. Again, it is difficult to maintain that latency does not increase with word length, as Santiago et al. do.

Santiago et al. observed that the complexity of the nucleus does not affect the production latency. The mean production latencies for the words with simple and complex nuclei were 273 and 265 ms, respectively, and did not differ statistically (see Figure 3). Santiago et al. take this null result (or “the shorter mean latencies”, as they call this 8-ms difference, p. 34) as evidence against a length account, but this conclusion is not warranted. The finding might simply mean that complex nuclei operate as simple segments at the phonological level (although they may be “complex” at

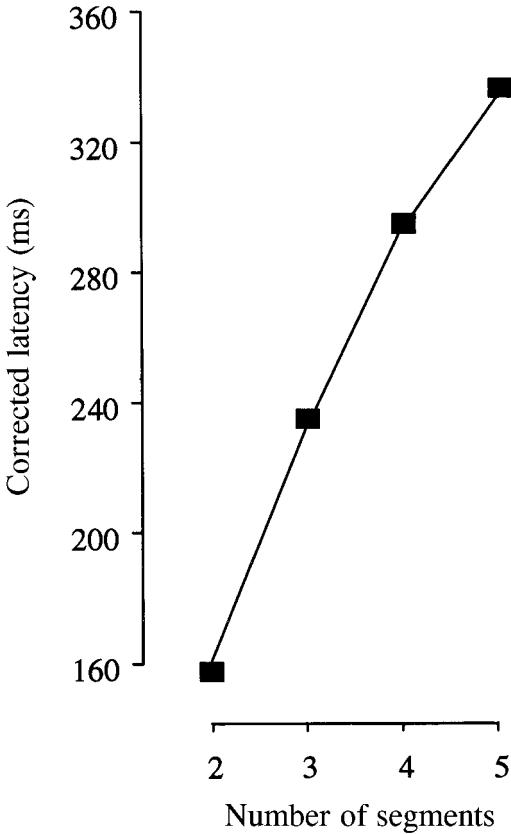


Figure 2. Corrected production latency (in ms) as a function of word length in terms of number of speech segments as measured by Santiago et al. (2000) in Experiment 2.

the phonetic level, cf. Roelofs, 1999).² That complex nuclei (i.e., diphthongs) act as single phonological segments has always been a working assumption in *WEAVER++*. I refer to the computer simulations reported in Roelofs (1996) and the *WEAVER++* lexical network shown in Figure 1 of that paper.

Coda complexity increased the latency by a non-significant 43 ms, which is numerically quite substantial (though Santiago et al. call it a “small” effect, p. 36) and clearly in the right direction for the length hypothesis (see Figure 3).³ As a comparison, the “large” (as they called it, p. 22) onset

² Under this assumption, there was no difference in word length between the words in the simple and the complex nucleus condition, $F(1, 20) = 1.11$, $MSE = 0.15$, $p = .304$.

³ There was a difference in word length between the words in the simple and the complex coda condition, $F(1,20) = 169.0$, $MSE = 0.04$, $p = .000$.

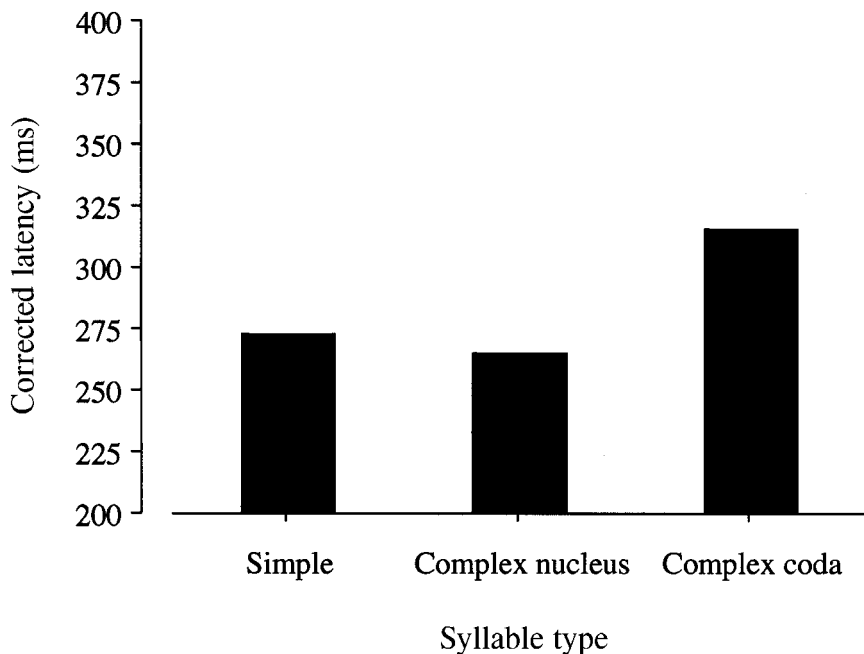


Figure 3. Corrected production latency (in ms) as a function of type of syllable as measured by Santiago et al. (2000) in Experiment 2.

complexity effect in Experiment 1 was 36 ms. Nevertheless, Santiago et al. maintain that the length account should be rejected because “it predicts that consonant clusters at the beginning of a word should delay vocal latency to the same extent as consonant clusters at non-initial parts of the word, contrary to present data” (p. 34). But as Santiago et al. admit, the failure to detect the 43-ms coda effect statistically might simply be due to insufficient power of the experiment (there were 12 items per condition, whereas the onset complexity effect of Experiment 1 was detected with 28 items per condition). Thus, the “absence” of an effect of coda complexity cannot be a serious argument against the length account, as Santiago et al. take it to be.

As another argument against the length account, Santiago et al. mention that the combined mean latency for the “CVVC and CVCC words” does not statistically differ from the mean latency for the “CVC words” – the difference is a non-significant 20 ms effect in the right direction. The CVVC/CVCC words are taken to be longer than the CVC words by Santiago et al. (pp. 34–35). However, “CVVC”, “CVCC”, and “CVC” are simply labels for the conditions in the experiment. The actual difference in word length between the conditions is, on average, only half (i.e., 0.5) a

segment. Given this small difference in word length (and the low power), the absence of a length effect does not come as a surprise and certainly does not challenge the length account. Similarly, the difference between the combined mean latencies of the “CCVVC and CCVCC words” and the “CCVC words” is not significant, but again, the actual difference in word length is only half a segment on average (0.4 of a segment, to be precise). To conclude, since the conditions differ by less than a segment, the absence of a length effect cannot be taken as a serious argument against the length account, as Santiago et al. take it to be.

DISCUSSION AND CONCLUSIONS

In summary, a re-analysis of the data of Santiago et al. shows that there is no support for effects of onset complexity and number of syllables when word length is taken into account. Instead, the syllable structure effects turn out to be word length effects. The observation that onset complexity and number of syllables do not play a role independent of word length poses difficulty to the NST. According to this theory, word length should have *no* effect, but my re-analysis of the data shows that length does play a critical role. In both experiments, production latency increases with word length. When one adjusts the statistical analyses of the production latencies for the effect of the uncontrolled word length, the variables onset complexity and number of syllables have *no* effect at all. Thus, the data fully agree with models such as WEAVER++ that predict effects of word length but not of onset complexity and number of syllables per se.

Santiago et al. refer to two other studies in the literature, in particular, Levelt and Wheeldon (1994) and Bachoud-Lévi, Dupoux, Cohen, and Mehler (1998), which failed to find length effects, and are therefore taken to be problematic for a length account of their onset complexity and number of syllables effects. They also mention findings from oral reading, but maintain that these data “can only be viewed as suggestive regarding production processes because reading tasks also involve perceptual processes” (p. 33). That is, any effects of syllable onset complexity and number of syllables could reflect the perception rather than the production side of a reading task. The aim of my comment was to address the rejection of models such as WEAVER++ by Santiago et al. on the basis of their own findings. Nevertheless, since this may leave the incorrect impression that I accept their arguments against the length account based on previous data, I end by briefly addressing the other arguments. I do not address the oral reading data, which Santiago et al. themselves already do not find very convincing.

First, it should be mentioned that a number of studies have obtained length effects on production latencies while keeping number of syllables

and syllable onset complexity constant (e.g., Roelofs, 1998; Roelofs & Meyer, 1998). So, it is not the case that there are no reliable effects of word length per se reported in the literature. Nevertheless, length effects are not always obtained. Santiago et al. point to an experiment by Levelt and Wheeldon (1994, Experiment 4), which failed to find a production latency difference for disyllabic words which differed in the length of the second syllable, 3 versus 5 segments. The problem with this argument is that it is based on a null result, which is always difficult to interpret. Experiment 4 of Levelt and Wheeldon (1994) was run as a control experiment just to show that a similar length difference in an earlier experiment could not be responsible for obtained effects. More importantly, there exists an interpretation compatible with WEAVER++ of this null result, which also may hold for the other recalcitrant finding referred to by Santiago et al., namely the null result for word length obtained by Bachoud-Lévi et al. (1998).

Left-to-right encoding with the phonological word as minimal unit for articulation initiation predicts that the production onset latencies should be larger for long (e.g., disyllabic) than for short (e.g., monosyllabic) words. However, Bachoud-Lévi et al. (1998) failed to find different production latencies for monosyllabic and disyllabic target words in picture naming experiments in French and English. They concluded that either word forms are not generated sequentially or that speakers start articulating before completing the planning of a phonological word. Either way, the theory of Levelt et al. (1999) would have to be changed. For example, WEAVER++ predicts a latency difference of 26 ms for the production of the monosyllabic and disyllabic words in the experiments of Meyer (1990, 1991) and Meyer and Schriefers (1991), contrary to the findings of Bachoud-Lévi et al.

However, in a replication of the experiments of Bachoud-Lévi et al. (1998) in Dutch, Meyer, Levelt, and Roelofs (in preparation) *did* obtain the length effect (of about 30 ms), but only when the materials were blocked for length. When monosyllabic and disyllabic names were mixed, as in the study of Bachoud-Lévi et al. (1998), the latencies did not differ significantly. Therefore, Meyer et al. assume, following Lupker, Brown, and Colombo (1997), that speakers can strategically control the criterion for the initiation of articulation. The criterion is optimal for each word type in blocked sets, but suboptimal for long and short words in mixed sets, which may conceal length effects. Van der Plaats and Van Galen (1990) observed the same “regression to the mean” for handwriting in mixed sets and made similar assumptions about the response criterion to account for their observations. WEAVER++ makes a distinction between *phonological encoding*, the process that generates the phonological word representations, and *phonetic encoding*, the subsequent process that

recovers stored phonetic syllable programs from memory for the syllables in the constructed phonological word. A possibility is that in mixed sets a phonological word representation is planned before articulation, but that articulation for disyllabic words starts after recovery of the phonetic program for the first syllable and that the second syllable program is accessed during the articulation of the first syllable.

Computer simulations showed that by a corresponding manipulation of the response criterion in WEAVER++, the difference between the latencies of monosyllables and disyllables disappeared, in agreement with the empirical findings of Bachoud-Lévi et al. (1998). Also, an early response criterion would explain why Levelt and Wheeldon (1994) failed to find an effect of the length of the second syllable (and why later studies failed to replicate their effect of second syllable frequency). Independent support for the view that a response criterion is set for the process that *follows* phonological encoding (i.e., phonetic encoding) comes from a study by Meyer and Schriefers (1991). They observed a stimulus onset asynchrony (SOA) difference for the effects of first and second syllable primes in mixed sets (suggesting sequential phonological encoding), but the magnitude of the auditory priming effects was equally large for both types of prime. This latter finding suggests that the response criterion applies after the phonological encoding of the second syllable of disyllabic words. If the criterion is set earlier (i.e., before the phonological encoding of the second syllable), different effect sizes for the first and second syllable primes are to be expected, contrary to the empirical findings.

Clearly, it is important to examine what exactly determines whether and when participants set a response criterion. In the study of Meyer, Levelt and Roelofs (in preparation), a latency difference (of about 30 ms) between monosyllabic and disyllabic words was obtained in *both* mixed and pure sets when participants were given feedback during the experiment about the speed of their performance. This suggests that with feedback, participants respond as fast as possible and do not set a response criterion. This also seems to have happened in the experiments of Santiago et al., where effects of length were obtained, as my re-analyses of their data showed. It is a topic for future research to determine when and how participants strategically control their response criterion, but this is irrelevant for now. The important point is that length effects have been obtained even with the types of materials used by Bachoud-Lévi et al. (1998).

To conclude, Santiago et al. (2000) claim to have obtained effects of onset complexity and number of syllables. These results are taken to support MacKay's (1987) Node Structure theory and are said to refute models such as WEAVER++ (Roelofs, 1997a) that predict effects of word length but not of onset complexity and number of syllables per se. In this

comment, I showed that a re-analysis of the data of Santiago et al. does not support their conclusions. Instead, the observed effects of onset complexity and number of syllables turn out to be length effects, in line with WEAVERTM and contrary to the Node Structure theory. In fact, the data of Santiago et al. present some of the best evidence that I know of for the claim that word production latency increases with word length.

Manuscript received March 2000

Revised manuscript accepted April 2000

REFERENCES

- Bachoud-Lévi, A.-C., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language*, *39*, 331–346.
- Levelt, W.J.M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, *42*, 1–22.
- Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38.
- Levelt, W.J.M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, *50*, 239–269.
- Lupker, S.J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 570–590.
- MacKay, D.G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer-Verlag.
- Meyer, A.S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, *29*, 524–545.
- Meyer, A.S. (1991). The time course of phonological encoding in language production: The phonological encoding inside a syllable. *Journal of Memory and Language*, *30*, 69–89.
- Meyer, A.S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 1146–1160.
- Meyer, A.S., Levelt, W.J.M., & Roelofs, A. (in preparation). Where is the length effect? Right here! A reply to Bachoud-Lévi et al. (1998).
- Roelofs, A. (1996). Serial order in planning the production of successive morphemes of a word. *Journal of Memory and Language*, *35*, 854–876.
- Roelofs, A. (1997a). The WEAVERTM model of word-form encoding in speech production. *Cognition*, *64*, 249–284.
- Roelofs, A. (1997b). Syllabification in speech production: Evaluation of WEAVERTM. *Language and Cognitive Processes*, *12*, 659–696.
- Roelofs, A. (1998). Rightward incrementality in encoding simple phrasal forms in speech production: Verb-particle combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 904–921.
- Roelofs, A. (1999). Phonological segments and features as planning units in speech production. *Language and Cognitive Processes*, *14*, 173–200.
- Roelofs, A., & Meyer, A.S. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 922–939.

- Santiago, J., & MacKay, D.G. (1999). Constraining production theories: Principled motivation, consistency, homunculi, underspecification, failed predictions, and contrary data. *Behavioral and Brain Sciences*, *21*, 55–56.
- Santiago, J., MacKay, D.G., Palma, A., & Rho, C. (2000). Sequential activation processes in producing words and syllables: Evidence from picture naming. *Language and Cognitive Processes*, *15*, 1–44.
- Van der Plaats, R.E., & Van Galen, G.P. (1990). Effects of spatial and motor demands in handwriting. *Journal of Motor Behavior*, *22*, 361–385.
- Winer, B.J., Brown, D.R., & Michels, K.M. (1991). *Statistical principles in experimental design*. New York: McGraw-Hill.