Monitoring the Inner Speech Code

Jane L. Morgan (J.L.Morgan.20@bham.ac.uk)
School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK

Linda R. Wheeldon (L.R.Wheeldon@bham.ac.uk)
School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK

Abstract

The aim of this paper is to expand and replicate the findings of Wheeldon and Levelt (1995). They employed an internal speech monitoring task which required Dutch speakers to monitor silently generated words for target syllable or phoneme sequences. On the basis of the obtained data several claims were made concerning the locus, time-course and nature of the internal speech code. The series of experiments reported here examined these predictions using English stimuli. In contrast to the Dutch study, no evidence of any reaction time advantage to syllable over nonsyllable strings was found. A phoneme monitoring experiment replicated the left-to-right pattern of results observed by Wheeldon and Levelt. In addition, a perception version of the task failed to replicate these effects suggesting that they were independent of the position of the target in the speech stream. Implications of the results in terms of the time course of phonological encoding are discussed.

Introduction

In this paper we report a series of experiments which investigate the time course of the generation of an abstract phonological representation during speech production. Current models of speech production postulate a process of phonological encoding which specifies a word's phonemic and prosodic properties. Following the selection of a word based on its semantic and syntactic specifications, it is proposed that a word's constituent phonemes and its metrical frame (e.g., syllable structure and stress pattern) are made available and are subsequently combined during a process termed segment-to-frame association (Levelt and Wheeldon, 1994; Roelofs, 1992, 1997). The output from this process is a syllabified phonological code which forms the input to phonetic encoding processes which retrieve syllabic gestural scores detailed enough to guide articulation. With regard to phonological encoding recent theories propose that the process of segment-to-frame association occurs in sequence from left-to-right across a word and a number of experimental findings exist which suggest that the beginning of a word is phonologically encoded before the end (Meyer, 1990, 1991; Meyer and Schriefers, 1991; Levelt, Roelofs and Meyer, 1999). However, there is a lack of data that would allow the temporal properties of such a process to be specified in any detail.

Our main aim was to replicate and extend the findings of Wheeldon and Levelt (1995) who developed an internal speech monitoring task in order to investigate the time course of the phonological encoding process. The phenomenon of inner speech is one which we all might experience, for instance whilst reading, writing, problem solving or whilst performing short term memory tasks. During such activities we often generate and monitor some form of internal speech code. Moreover, there is evidence to suggest that the generation of this code entails the same mechanisms which are employed in normal speech production (Dell and Repka, 1992; Levelt, 1989; Motley, Camden and Baars, 1982). Wheeldon and Levelt (1995) designed a series of experiments which exploited this ability to monitor one's own inner speech. They adapted a methodology traditionally employed in the field of speech perception in which subjects monitor spoken words for target syllables and phonemes (Cutler, Mehler, Norris and Segui, 1986; Mehler, Dommergues, Frauenfelder and Segui, 1981; Zwisterlood, Schriefers, Lahiri and van Donselaar, 1993). Native Dutch speakers were required to silently generate Dutch translations (e.g., kado) of English prompt words (e.g., gift) and to monitor the Dutch carrier words for a prespecified target sequence (e.g., /k/ or /ka/).

The main findings of these experiments were interpreted as indicating that the code being generated by subjects during the monitoring tasks was an abstract phonological one. This was based on the demonstration that subjects' ability to perform the monitoring task was unimpaired by the addition of a secondary articulatory suppression task. It was also argued that the code could not be phonetic as no relationship was found between the pattern of monitoring latencies and the timing of the same targets in overt speech production. Furthermore it was claimed that subjects were not basing their responses on the initial availability of the word's phonemes but rather on the output of the segment-to-frame association process as it was found that the code being generated was syllabified. When subjects were required to monitor bisyllabic words for target strings which either did or did not correspond to the first syllable in a word (targets /ma/ and /maag/ in the carrier words *ma.ger* thin, and *maag.den* - virgin) reaction times were faster to the target sequence which was a syllable in the carrier word irrespective of target size. Importantly this syllabification process was argued to be on-line and not merely involve the downloading of stored syllable units.

Subjects were also instructed to monitor CVC.CVC carrier words (e.g., *lif.ter*) for the four constituent consonants and it was seen that the time taken to detect each segment increased in a left-to-right manner. In addition, a 50ms difference was observed between the monitoring latencies for the two medial phonemes which flanked the syllable boundary. This difference was similar in size to that found in monitoring latencies for the first syllable consonants which are separated by an intervening vowel. This would concur with the notion that within syllables phonemes are assigned to their frame in a left-to-right manner. The authors argued that the syllable boundary effect suggests that this assignment process is initiated for the first syllable and the encoding of the second syllable is held up until it is completed. It was also observed that there was a significant difference in the time taken to monitor for the two phonemes in the first syllable but this difference was not seen for the phonemes in final syllable. This was taken as evidence that even though the encoding of the second syllable of the word is delayed the constituent phonemes are still made available. This means that when assignment to the syllable frame can begin it occurs much faster than that of the previous syllable.

While the results of the Wheeldon and Levelt (1995) are intriguing a couple of aspects of their data are problematic. The first problem concerns the locus of the observed syllable monitoring effect. As the authors acknowledge, it is possible that the effects being observed, especially those concerning the structural properties of inner speech, might be a function of the monitor itself rather than the actual speech code being monitored. A number of monitoring devices have been proposed which fall into two broad categories: production and perception based monitors. It is possible that speakers can monitor their speech at every level of the speech production process (from conceptualisation onwards) as proposed in models such as Laver (1980). However, this implies much duplication between the monitor and the actual process itself. A more parsimonious account holds that it is only the output of phonological processing which can be monitored (Motley et al, 1982). Although this model overcomes the duplication issue it still does not satisfactorily explain how speakers are able to monitor their speech on a variety of other linguistic levels (i.e. for semantic, syntactic, social appropriateness, Levelt, 1983; 1989). In order to address the shortcomings of such production based models, Levelt (1989) outlines an device which uses as its input the internal speech code but feeds it through an internal loop into the speech comprehension system. Put simply, Levelt argues that prearticulatory speech is monitored and edited as if it where the speech of an external speaker. If this is the case, then it follows that any observed effects arising from Wheeldon and Levelt's monitoring task could, at least in part, be due to the parsing processes of the speech perception mechanisms.

Zwisterlood et al (1993) observed a syllable match effect when subjects were required to monitor auditorily presented Dutch words. It seems, therefore, that Dutch speakers make use of syllable units during their monitoring of an incoming speech stream. However, this is not universally true of all languages. It was, therefore, thought worthy to investigate whether Wheeldon and Levelt's findings can be replicated in a language which does not appear to employ the syllable in the same way during monitoring. English is the ideal candidate. It is a language which contains many instances of words which are ambisyllabic, has an irregular stress pattern and contains a large range of syllabic structures. This means that employing a syllable-based segmentation strategy would be unproductive (Cutler, 1997; Cutler et al., 1986; Cutler and Norris, 1988; Norris and Cutler, 1985). Indeed, there has been no evidence of a syllable effect using a monitoring task in the perception of English (Cutler, 1997; Cutler et al., 1986; Bradley, Sanchez-Casas and Garcia-Albea, 1993). By contrast there is much evidence to suggest that the syllable is a salient unit in the production of English (Ferrand, Segui and Humphreys, 1997; Sevald, Dell and Cole, 1995). If a syllable effect can be demonstrated in the internal production of English it would add weight to the notion that the properties observed are ones of the production based speech code and not features of the comprehension system by which it is monitored.

The second issue concerns the predictions made regarding the time course of phonological encoding. These were based chiefly on the results of the phoneme monitoring task which has some methodological shortcomings. The left-to-right effect was not robust in that it only reached significance in the subjects analysis of the data and not the items analysis. In addition, the significant monitoring difference which was found between the two phonemes on either side of the syllable boundary was confounded by the fact that for the majority of stimuli this also served as a morphological boundary. As a consequence there is a need not only to replicate this pattern of results but also to employ a more stringent set of stimuli.

Finally, it was thought that a fruitful way to extend the Dutch study would be to directly compare reaction times on the inner speech monitoring task with an identical perceptual version. It still cannot be argued unequivocally from the Wheeldon and Levelt data that the pattern of results obtained from the phoneme monitoring study is one which is specific to speech production. In particular the speeding up of the encoding of the final syllable could be a feature of speech perception mechanisms. For instance, once the uniqueness point of a word has been reached processing could progress at a faster rate. For this reason it is important to establish how closely the monitoring latencies observed in the inner speech experiment correspond to those obtained using a perception task.

The Task

It was not possible to exactly replicate the methodology employed by Wheeldon and Levelt (1995) as they relied upon a population of subjects who were comfortably bilingual and unfortunately such subjects were not available to us. Instead a task was employed which made use of word association pairs. Specifically the word forms to be monitored were elicited using a semantically related prompt words (e.g., prompt *baboon*, carrier word *monkey*). The prompt words were chosen carefully to avoid any potential priming effects due to phonological relationships with the associated word. In an attempt to encourage the use of the internal speech code rather that any visual or graphemic representations the target sequences and prompt words were always presented auditorily and the pre-experimental training which the subject undertook in order to learn the word pairs was given verbally.

Experimental trials were structured as follows. First an auditory description of the target sequence was presented (e.g., /mon/). This was followed by the auditory presentation of a prompt word (e.g., *rain*). Subjects made a push button yes/no response depending upon whether the carrier word (e.g., *monsoon*) contained the target string. Reaction times and errors were recorded. Responses were measured from the onset of the prompt word to the subject's response. All reported statistical analyses were conducted treating subjects (t_1 and F_1) and items (t_2 and F_2) as random factors.

Investigating the Syllable Effect

Experiment 1 - Initial Syllable

Method This experiment investigated whether English speaking subjects would detect strings which corresponded to the first syllable in a word faster than if the sequence constituted more than a syllable. Carrier words were chosen with CVC.C (or VC.C) syllable boundaries. Words were selected for which syllabification was unambiguous according to the phonotactic rules of the language. For instance, the word *tempest* is syllabified *tem.pest* , not *temp.est* because according to the maximisation of onset principle the second syllable onset will attract the /p/ consonant but not the /mp/ cluster which is an illegal syllable onset in English. The target sequences to be monitored for were the initial CVC/VC (syllable - e.g., /tem/) and CVCC/VCC (nonsyllable - e.g., /temp/).

Results and Discussion The mean monitoring latencies for the two target sequence types are given in Table 1. It can be seen that subjects were quicker at detecting a nonsyllable string than a syllable string. A related samples t-test was performed and confirmed that this difference in means was significant in the subjects analysis only, t_1 (23) = 3.0, Standard Error = 16, $p < .01$. The same analyses were conducted using the error rates and no significant differences were observed between the two conditions for either analysis.

Table 1: The Mean Latencies (in ms) for the Two Strings Employed in the Syllable Monitoring Experiments. Mean Percentage Errors are Presented in Brackets.

Experiment	Syllable	Nonsyllable	Difference
(1) Initial CVC	1144 (4.2)	1096 (4.2)	48
(2) Internal CVC	1431 (9.4)	1350 (6.8)	82
(3) Initial CV	1308 (5.5)	1239 (5.5)	70

This experiment, therefore, yielded no evidence of a syllable match effect. Instead there was an insignificant tendency for the longer CVCC target to be monitored for faster than the CVC syllable target. It is possible that a priming effect swamped any syllable effect which might be present; simply the more of the word that is heard the quicker it can be processed. As a consequence the next experiment was designed in order to dissociate the potential syllable effect and the influence of priming by requiring subjects to monitor the carrier word for word internal syllables. If the sequence to be monitored for is no longer coming from the initial portion of the carrier word then hearing the string beforehand should not prime its generation (Meyer, 1990; 1991).

Experiment 2 - Internal Syllable

Method The method was exactly the same as employed in the previous study with the exception that the target strings occurred internally in the carrier word. The syllable sequence to be monitored for was of CVC (and in one instance, CCVC) structure and the nonsyllable sequence CVCC (or CCVCC). For example, the carrier word *romantic* would be preceded by the target strings /man/ (syllable) and /mant/ (nonsyllable).

Results and Discussion As can be seen in Table 1 the time taken to monitor for the nonsyllable string was again faster than for the string which corresponded to a syllable in the carrier word. This difference was significant according to both the subject, t_1 (23) - 4.0, SError = 20, $p < .01$ and the item t_2 (15) = -2.2, SError = 40, $p < .05$ analysis. No significant difference was found between the number of errors made in each condition. Therefore, even when the position of the target sequence is changed to an internal sequence the effect remains constant in that the longer string produces the faster reaction times. These data contradict the notion that any syllable match effect is being masked by a priming effect.

From the data of these first two experiments it must be concluded that there is no evidence that subjects' monitoring latencies are faster when the sequence to be detected in their internal generations corresponds to a syllable in that word. However, there still remains the possibility that the observed effect is a feature of the stimuli being utilised. As outlined in the Introduction the English language is notoriously ambisyllabic and one factor which influences syllabification in English is stress. Treiman and Zukowski (1990) using an oral syllable repetition task observed that for words with a medial consonant cluster (such as those used in Experiment 1 and 2 - for example, tem.pest) the second consonant (so in this example the /p/) is likely to be attached to the first stressed syllable in addition to the onset of the following unstressed syllable; in other words become ambisyllabic. This should not be overstated as the task used did not reflect normal speech production processes and the effect was seen for only certain types of consonant clusters. However, as this ambisyllabic predisposition is one which holds for most of the stimuli described so far a final syllable monitoring experiment was conducted which employed second syllable stress words and avoided medial consonant clusters.

Experiment 3 - Initial Syllable with Noninitial Stress

Method The same design as employed in the previous two experiments was replicated but words of CV.CVC or (V.CVC) structure acted as the carrier words (e.g., *pla.toon*). Gussenhoven (1986) and Treiman and Danis (1988) have identified certain rules which make ambisyllabic intervocalic consonants less prevalent. Specifically, a consonant is more likely to be syllabified with the following vowel when that vowel is stressed; in addition obstruent consonants are more often grouped with the following vowel than sonorant consonants. Following this second syllable stress words were used and it was ensured that the second syllable onset was always a stop consonant or fricative. The syllable sequence which was to be monitored for was of CV structure (e.g., /pla/) and the nonsyllable sequence was of CVC composition (e.g., /plat/).

Results and Discussion As in Experiments 1 and 2 the nonsyllable sequence produced faster monitoring latencies than the syllable sequence. The respective means are detailed in Table 1. This difference was significant by subjects analysis only, t_1 (23) = 3.3, SError = 21, p < .01. The two conditions did not significantly differ in the number of errors which were observed.

Therefore, yet again there is no evidence to suggest that the syllable sequence has a facilitative effect on monitoring latencies. Indeed this is the indisputable finding from the entire series of experiments. It has been consistently found that the trend in subjects' reaction times goes against that which would be predicted by the syllable effect hypothesis. This is a very similar pattern of results as seen for English in the perception studies (Cutler et al 1986; Bradley et al 1993) which casts doubt on the conclusion drawn by Wheeldon and Levelt, namely that their syllable effect reflects properties of the production code. The lack of a syllable effect in the present studies seems to verify the internal loop theory of monitoring (Levelt, 1989) and specifically that it is the properties of the speech comprehension system which are reflected in the syllable monitoring task.

The next two experiments to be reported focused on the monitoring of phonemes in bisyllabic words. This was to establish whether the claims made by Wheeldon and Levelt regarding the time course of phonological encoding could be repeated.

The Time Course of Phonological Encoding

Experiment 4 (a) and (b) - Phoneme Monitoring

Method - Experiment 4 (a) The task employed was similar to that described for the syllable monitoring experiments. The carrier words used were of CVC.CVC structure with a clear, phonotactically correct syllable boundary (e.g., *lit.mus*). The majority of the words were monomorphemic and each contained four different consonants which served as the target phonemes which were to be detected during the task. In this manner the four target positions were situated at the first syllable onset (C1); first syllable offset (C2); final syllable onset (C3); and final syllable offset (C4). All the intervening vowels (with exception of three) were short.

Table 2: The Mean Monitoring Latencies (in ms) for the Four Target Positions In Experiments 4 (a) and (b). The Mean Percentage Errors are Detailed in Parenthesis.

	Target Type			
Experiment	C1	C2	C3	C4
Internal Production	1392 (3.5)	1501 (6.0)	1564 (4.8)	1589 (5.0)
External Perception	622 (1.6)	886 (1.6)	901 (0.9)	1028 (2.3)

Monitoring trials were grouped into lists for a given target phoneme. Such lists comprised of between four and twelve prompt words. Subjects heard a description of the target phoneme followed by a series of prompt words. For each of these the subject was required to decide whether the corresponding carrier word contained the sound described to them. If the carrier word did contain the phoneme a, 'yes' response was required, if it did not, no response was necessary.

Results and Discussion As can be seen in Table 2 the mean latencies increase as a function of their position in the word in a left-to-right manner. The differences in monitoring latencies across consonant positions are given in Table 3. The difference between the initial syllable onset and offset is the greatest, followed by that between the consonants which flank the syllable boundary with the difference between the constituent phonemes of the final syllable being the smallest. ANOVAs were performed on the data and a significant difference between the four target positions was demonstrated, F_1 (3, 108) = 30.5, MSError = 9540, $p < .01$, F_2 (3, 51) = 19.8, MSError = 7350, $p < .01$. Newman-Keuls pairwise comparisons yielded a significant difference between C1 and C2 and between C2 and C3. The difference between C3 and C4 was not significant. Identical ANOVAs were repeated using the error rates which yielded no significant effects.

This pattern of results is identical to that found by Wheeldon and Levelt (1995). In this way the effects demonstrated by the phoneme monitoring methodology appear to be more robust across languages than those involving syllable monitoring.

Method - Experiment 4 (b) This was an attempt to establish whether the pattern of latencies observed on the production task bore any resemblance to the actual position of the target phonemes in the speech stream. The experiment was repeated but instead of requiring subjects to generate the carrier words internally they were presented to them auditorily.

Table 3: A Comparison of the Difference in Monitoring Latencies (in ms) between the Four Target Positions in Experiments 4 (a) and (b). Percentage Error Scores are Given in Brackets.

	Difference	Experiment Internal Production	 External Perception
(0)	C1-C2	109 (2.5)	264
	C2-C3	63 (1.2)	16 (0.7)
	C3-C4	25 (0.2)	127 (1.4)

Results and Discussion The data for this perception experiment is also shown in Tables 2 and 3. Once again the mean monitoring latencies for the four experimental conditions increases the further along in the word the phoneme is positioned. ANOVAs demonstrated that the difference between these conditions was significant, F_1 (3, 69) = 118.4, MSError = 5893, $p < .01$ and by items, F_2 (3, 51) = 44.8, MSError = 11770, $p < .01$. Pairwise comparisons showed that the difference between the two syllables onsets and offset were significant. In contrast, the difference between the two phonemes which flanked the syllable boundary was not. This reflects a different pattern of monitoring latencies to the ones observed in the production experiment but is consistent with the position of the target segments in the speech stream in that the syllable onsets and offsets are temporally separated by a vowel whereas C2 and C3 are not. This would suggest that the code being monitored in the internal speech task is an abstract one.

General Discussion

In summary, this paper reports a series of experiments which required subjects to monitor their internal productions of English words for target syllables or phonemes. Experiments 1-3 show that there is no evidence of a syllable match effect in English. However, when subjects were required to monitor bisyllabic CVC.CVC words for their constituent phonemes a clear left-to-right effect was observed. In addition, the time taken to detect the phonemes on either side of the syllable boundary was seen to differ and a significant difference was seen between the monitoring latencies for the first syllable onset and offset which was not repeated in the final syllable. These findings relate to those of Wheeldon and Levelt (1995) as follows.

Regarding the locus of the syllable monitoring effect, our results seem to reflect characteristics of the perception rather than the production mechanisms. This supports the claim that the prearticulatory speech code is monitored by the speech comprehension system via an internal loop and that this architecture is not sensitive to syllable structure for English (Cutler, 1997; Cutler, et al., 1986; Cutler and Norris, 1988, Norris and Cutler, 1988).

In contrast, however, the claims regarding the time course of the phoneme monitoring task hold true in that the above results mirror those obtained in the Dutch study. It has again been shown that speech encoding runs in a left-to-right manner and that the code is syllabified. Importantly, the syllable boundary effect was replicated with words which were not morphologically complex supporting the idea that the syllabification of subsequent syllables is held up until phonemes have been assigned to the initial syllable. These findings compliment current models of speech production such as proposed by Levelt, Roelofs and Meyer (1999) and the computer-based WEAVER model (Roelofs, 1992; 1997) which propose that phonological encoding runs in a strictly serial order and that the syllabification of an utterance is generated on-line in accordance with the rules of the language.

Finally, it was found that when the phoneme monitoring task was repeated using external speech a completely different pattern of results was seen. Specifically, it was seen that the speeding up of encoding in the final syllable which was observed in the inner speech task was not found in the perception version. This confirms that the code being monitored during the inner speech task is phonological and not encumbered by phonetic or articulatory specifications.

In this way the data presented in this paper has been able to resolve some of the problematic aspects of the Wheeldon and Levelt studies. The lack of a syllable matching effect in the above data allows certain aspects of the Dutch data to be attributed to a comprehension not production-based monitor. It has been shown that the left-to-right pattern is robust across different languages and different word eliciting tasks. Finally, it has been confirmed that the time course of internal speech monitoring is different to external speech in theoretically interesting ways. Thus, the internal speech monitoring task can be seen to be a valid methodology for the investigation of the time course of phonological encoding during speech production.

Acknowledgements

This research was conducted by the first author and funded by means of an ESRC research grant awarded to the second author. The authors wish to thank Jan Zandhuis for the design and programming of the experimental set-up.

References

Bradley, D. C., Sanchez-Casas, R. M., & Garcia-Albea, J.E. (1993). The status of the syllable in the perception of Spanish and English. *Language and Cognitive Processes, 8*, 197-233.

Cutler, A. (1997). The syllable's role in the segmentation of stress languages. *Language and Cognitive Processes, 12*, 839-845.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language, 25*, 385-400.

Cutler, A., & Norris, D. (1988). The Role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 113-121.

Dell, G. S., & Repka, R. J. (1992). Errors in inner speech. In B. J. Baars (Ed) *Experimental Slips and Human Error: Exploring the Architecture of Volition.* New York: Plenum Press.

Ferrand, L., Segui, J., & Humphreys, G. W. (1997). The syllable's role in word naming. *Memory and Cognition, 25*, 458-470.
Gussenhoven, C. (1986). English plosive allophones and ambisyllabicity. *Gramma, 10*, 119-141.
Laver, J. D. M. (1980). Monitoring systems in the neurolinguistic control of speech production. In V. A. Fomkin (Ed) *Errors in Linguistic Performance.* New York: Academic Press.
Levelt, W. J. M. (1983). Monitoring and self repair. *Cognition, 14*, 41-104.
Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioural and Brain Sciences, 22*, 1-75.
Levelt, W. J. M., & Wheeldon, L. R. (1994). Do speakers have access to a mental syllabary? *Cognition, 50*, 239-269.
Mehler, J., Dommergues, J., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behaviour, 20*, 298-305.
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: 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.
Motley, M. T. , Camden, C. T., & Baars, B. J. (1982). Covert formulation of anomalies in speech production: Evidence from experimentally elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behaviour, 21*, 578-594.
Norris, D., & Cutler, A. (1988). The relative accessibility of phonemes and syllables. *Perception and Psychophysics, 43*, 541-550.
Roelofs, A. (1992). A spread-activation theory of lemma retrieval in speaking. *Cognition, 42*, 107-142.
Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition, 64*, 249-284.
Sevald, C. A., Dell, G. S., & Cole, J. S. (1995). Syllable structure in speech production: Are syllables chunks or schemas? *Journal of Memory and Language, 34,* 807-820.
Treiman, R., & Danis, C. (1988). Syllabification of intervocalic consonants. *Journal of Memory and Language, 27*, 87-104.
Treiman, R., & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language, 29*, 66-85.
Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language, 34*, 311-334.
Zwisterlood, P., Schriefers, H., Lahiri, A., & van Donselaar, W. (1993). The role of syllables in the perception of spoken Dutch. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19*, 200-271.