

Selective attention and speech errors: Feature migration in time

Nazbanou Nozari (nnozari2@illinois.edu)

Beckman Institute, University of Illinois at Urbana-Champaign,
405 N Matthews Ave, Urbana, IL 61801, USA

Gary S. Dell (gdell@cyrus.psych.illinois.edu)

Beckman Institute, University of Illinois at Urbana-Champaign,
405 N Matthews Ave, Urbana, IL 61801, USA

Abstract

Studying speech errors has revealed much about the language production system. We believe it can also be used to investigate the interaction between the speech production and other systems, such as the executive system. This paper studies the effect of focusing attention on the production of a single word in a sequence. We present three experiments in which participants, while reciting four-word tongue twisters, allocate their attention to a particular word either by instruction to avoid errors on that word, to stress the word, or reversely, to recite that word silently instead of overtly. The results of all three experiments consistently show that focusing attention on one word causes a higher error rate on the other words in the sequence.

Keywords: Speech errors; Attention.

Introduction

To avoid mistakes, you must “pay attention”. That is what every parent, teacher, or coach will tell you. Attending to something, does in fact have measurable effects. Object features, such as color, can freely move to other objects if attention is not focused on the reported object at the time of its perception (e.g. Treisman & Gelade, 1980). For example, participants who were briefly exposed to an array of letters containing a green X and a red O, might mistakenly report having seen a red X, if the attention is diverted by some means. However, when the to-be-reported object’s location is precued, thus selectively focusing attention on that object, such “conjunction errors” decrease significantly, even when the overall difficulty of the task matches cases in which the object was not attended to (Treisman & Schmidt, 1982).

An equivalent of feature migration is found in reading (e.g. Allport, 1977; McClelland & Mozer, 1986; Mozer, 1983; Shallice & McGill, 1978; Shallice & Warrington, 1977), which in its extreme form manifests as attentional dyslexia (e.g. Davis & Colthart, 2002; Friedmann, Kerbel, & Shvimer, 2010; Humphreys & Mayall, 2001; Mayall & Humphreys, 2002; Saffran & Coslett, 1996), a disorder in which the high rate of letter migration between written words makes reading very difficult to impossible. The problem can be easily remedied by the use of a moving window, about the size of a single word, which provides a narrow focus for attention.

Speech production also suffers from “feature migration”, but migration in time, rather than migration in space. Parts of words (e.g. phonemes) move between the planned words

(e.g. Dell & Reich, 1981; Fromkin, 1971), giving rise to speech errors such as “dig booz” instead of “big dues”. Uncontrollable migration of phonological constituents might in fact be the underlying cause of the “naming-in-context” disorder in which aphasic patients cannot give the names of a sequence of pictured objects (e.g. Schwartz & Hodgson, 2002). Similar to attentional dyslexia, the problem is easily remedied by narrowing down attention to a single picture at a time by the means of removing or covering the other pictures.

In summary, attention seems to play a crucial role in both perception and production of language. While the former has been studied systematically, the latter has received little scrutiny. In the same vein, models of spoken speech errors have featured prominently in cognitive science for 25 years (e.g. Dell, 1986; Nozari et al., 2010), but no speech-error model takes any note of attention. The first step for implementation of such mechanisms is to gather empirical data on how the process of lexical access is affected by tuning attention to a word in a sequence. It is thus the goal of this paper to answer a basic question: If attention is focused on one word in a sequence of words, what happens to the other words? Does attending to a word increase the overall accuracy of the speaker or does it come at the cost of other words?

We present three experiments, in which we elicit speech errors by having participants recite a series of four-word tongue-twisters, with one (or none) of the words being attended. In Experiment 1, participants are directly instructed to avoid mistakes on a particular word. Experiment 2 aims to elicit an overt manifestation of focusing attention by asking the participants to *stress* a certain word among the four words. Finally, Experiment 3 asks participants to *mouth* (silently recite) one of the four words while the other words are spoken aloud normally. Thus Experiment 3 requires attention to a particular word, but the physical manifestation of that attention is the opposite of the stress instruction given in Experiment 2.

Experiment 1

Methods

Participants: 32 undergraduate students at the University of Illinois at Urbana-Champaign participated for course credit.

Materials: 32 pairs of tongue-twisters (total = 64) were borrowed from Oppenheim and Dell (2008). Each tongue-twister consisted of four words with the same vowels and an XYYX onset pattern. In half of the tongue-twisters the onsets were dissimilar (e.g., “just rum rug jump”). Each dissimilar tongue-twister was paired with another that shared its middle two words, but had one onset consonant changed so that the onsets were phonologically similar (e.g., “lust rum rug lump”). Four lists, each containing sixteen tongue-twisters, were constructed, two of which were used in the experimental and the other two, in the control condition. One half of each pair always appeared in the experimental condition and the other half in the control condition, such that a phonologically similar tongue-twister was half the time in the experimental and half the time in the control condition. Each list contained an equal number of phonologically similar and dissimilar tongue-twisters.

In the experimental condition, one of the four words in the tongue twister was printed in bold and was underlined (e.g., “just rum **rug** jump”). Each of the four words was so marked with equal probability across participants. In the control condition, no word was bold or underlined.

Procedure: Participants completed four test blocks, each of which contained one of the four lists of tongue-twisters. Half the subjects completed the experimental blocks first, and the other half, the control blocks first. The experiment was run in MATLAB. In the control condition, a single trial consisted, first, of a familiarization display, where the four words appeared on a computer screen and the participant was asked to enunciate each word carefully. Necessary corrections were made by the experimenter until the participant successfully pronounced all four words. Next, a rehearsal phase began, in which the same four words appeared on the screen and a metronome played at 2 beats/sec. For four beats, the participant only listened. At the beginning of the fifth beat, a red dot appeared under each word, and moved to the next word at the pace of the metronome. Participants were to read the words aloud as the red dot moved along. They then listened for the next four beats (without the red dot) and began reciting on the fifth as the red dot reappeared. Four rehearsal cycles were completed in each trial. No errors were coded at this phase. If the participant repeatedly made an error, the experimenter corrected the error and had them repeat the rehearsal phase of that trial from the beginning (this was seldom required). Finally, the test phase began. Like the rehearsal phase, there were four cycles, each containing a silent part in which the participants just listened to the metronome for four beats and a speaking-aloud part, in which they had to say the words aloud at the pace of the metronome, and along with the visual cue (the red dot). There were two differences between the rehearsal and the test phase though: first, the pace of the metronome was higher in the test phase (3 beats/sec), and second, the words were no longer printed in the center of the screen. In the silent part of each cycle, the words appeared in a small font in gray color at the top of the

screen. Participants were encouraged to repeat the words from memory, but were allowed to consult them if they could not remember them. In the speaking aloud part of the cycle, the words disappeared and the red dot appeared at the place previously occupied by the printed words in the center of the screen in the rehearsal phase. As the dot moved from position 1 to 2 to 3 to 4, along with the metronome beat, the participant had to say the words aloud. They were told to try to repeat the words as accurately as possible on the metronome beat.

After receiving the instructions, participants completed a short practice block of eight trials. If the experimenter judged their performance with regard to following the general instructions to be unsatisfactory at the end of the practice block, they were run through that block again.

The experimental condition was very similar, with the exception that one of the words was bold and underlined as soon as the participants got to the rehearsal phase. They were instructed to try to be accurate overall, but to especially avoid making errors on the underlined word or else they would hear an annoying buzz (the experimenter who monitored their performance used the buzzer from the game Taboo). The relevant instructions were presented right before the participants started this condition and a practice block comprised of eight trials preceded the test blocks. Similar to the control condition, participants did not start the test blocks unless satisfactory performance was reached on the practice block.

During the study, an experimenter (the first author or a trained assistant) sat next to the participants. In the experimental condition, s/he buzzed the participants immediately if they made an error on the underlined word in the test phase. Responses were digitally recorded for offline transcription. Both the first author and the assistant transcribed errors from the recordings. The original agreement was 76%. In cases where the two transcriptions differed, the recording was replayed, and if the transcribers still disagreed, the utterance was not coded as an error.

Results and Discussion

One participant was excluded from this experiment for failing to follow instruction. There were thus a total of 15872 spoken words from the remaining 31 participants, each of which represents an opportunity for error. For participants who did follow the instructions, if they refrained from responding on 50% or more of the words in a trial during the test phase, that trial was excluded. After enforcing this criterion, data from four trials (64 error opportunities), all belonging to the experimental condition, were eliminated.

Errors consisted of substitution, addition, or deletion of phonemes, or, less often, substitution of entire words. 82% of all errors were within-trial migrations, meaning that a phoneme had moved from one word to another in the same trial, or alternatively, a word had moved to a different position in the trial. The remaining 18% of the errors were

outside intrusions (e.g. "just *rub* rug jump" for the target "just rum rug jump"). Of the within-trial migrations, 29% were word substitutions. This, however, might be an overestimation, because coding word substitutions was prioritized over phoneme substitutions. For example, the first *rug* in "just *rug* rug jump" for the target "just rum rug jump" was coded as a word substitution (there was an identical word in the target that now appeared in the wrong place in the error), although theoretically it could have been a phoneme substitution.

In order to compare the experimental to the control condition, we assigned what we called "fake" attended status to words in the control condition. To this end, for each half of a tongue-twister pair in the experimental condition, we found the other half in the control condition and assigned the word in the same position to the fake attended status. In other words, we pretended that it had been underlined for other purposes of analysis. For example, if "just rum rug jump" was in the experimental condition, "just rum rug lump" was located in the control condition and "rug" was given the fake role of the attended word. From now on, we use the term "attended and unattended" for both experimental and control conditions, although in reality, there was no difference between the four words in the control condition.

A total of 804 errors (149 on the attended and 655 on the unattended words) were made in the experimental condition, and 716 errors (192 on the attended and 524 on the unattended words) in the control condition. Since the total opportunities for making errors differed, each number was turned into a proportion by dividing it by the total number of error opportunities a participant had in that condition for that word-status. Figure 1a shows these proportions. There were significantly more errors on the unattended than the attended words in the experimental condition ($t(30) = -3.63$; $p < .001$). The difference between the error rates on the (fake) attended and unattended words in the control condition was, however, not significant. This shows that the experimental manipulation did have the desired effect. To examine how focusing attention on one word affects the processing of the unattended words, we performed hierarchical binomial multiple regression in the statistical R package, using a mixed model which treated subjects and items as random effects. We also tested the need for a randomly varying fixed effect by both subjects and items. Random slopes for each predictor were added to the model one by one, in a fixed order. We compared the model with the newly added term to the preceding model by performing a Chi-squared test of the deviance in model log likelihood. If this was not significant, we dropped the term from the model and continued adding other random slopes. The final model consisted of condition (experimental vs. control), word-status (attended vs. unattended) and the interaction of the two as fixed effects, and subject and item random intercepts, as well as random slopes of condition for subjects and word-status for items as random effects.

There were more errors overall in the experimental condition, although the difference did not reach significance ($z = 1.65$, $p = .10$). Crucially, the interaction between condition and word-status was significant ($z = -3.43$, $p < .001$). Contrast-testing against a baseline of averaged (fake) attended and unattended words in the control condition revealed that there were significantly fewer errors on the attended words compared to the baseline ($z = 3.063$; $p = .002$), and significantly more errors on the unattended words compared to the baseline ($z = -2.543$; $p = .010$).

The results of experiment 1 suggest that focusing on one word is likely to hurt the overall performance by causing more errors on the other words in the sequence. It is, however, unclear what strategies the participants used in order to avoid mistakes on the attended word. In Experiment 2 we asked them to focus on the word by overtly "stressing" it.

Experiment 2

Methods

Participants: 32 undergraduate students at the University of Illinois at Urbana-Champaign, who had not taken part in experiment 1, participated in the study for course credit.

Materials: The materials from experiment 1 were used.

Procedure: The procedure was the same as for experiment 1, except for the instructions given for the experimental condition. Participants were told to stress the bold and underlined word when they pronounced it. An example was given: "John shot two ducks", and the experimenter showed the participants how each word can be given emphatic stress (modeled with the pitch accent L+H*) if the speaker meant to emphasize a certain aspects of the message.

Results and Discussion

The exclusion criteria of Experiment 1 were enforced here as well. In addition, if the participant could not follow the "stressing" instructions properly after completing the practice block twice, they were excluded. Two participants were thus eliminated. For the remainder of the participants, recitations during the test phase in which they did not stress the underlined word (as judged by the two independent transcribers) were thrown out. Overall, 336 error opportunities were excluded for the remaining 30 participants, all belonging to the experimental condition.

The same procedure as Experiment 1 was repeated to assign fake stressed status to words in the control condition. 1023 errors (189 on the stressed and 834 on the unstressed words) were made in the experimental, and 943 errors (258 on the stressed and 685 on the unstressed words) were made in the control condition. Figure 1b shows the proportion of errors for each word-status in each condition. Similar to

Experiment 1, there were significantly more errors on the unstressed, compared to the stressed words in the experimental condition ($t(29) = -3.61$; $p < .001$), but no significant difference between the stressed and unstressed words in the control condition. A similar type of hierarchical logistic regression model was used, and the same procedure described in Experiment 1 was repeated to decide which random effects to include. The final model consisted of condition (experimental vs. control), word-status (stressed vs. unstressed) and the interaction of the two as fixed effects, and subject and item random intercepts, as well as random slope of condition for both subjects and items, and random slope of word-status for items as random effects. The findings were the same as Experiment 1. Although there were overall more errors in the experimental condition, this difference did not reach significance ($z = 1.48$, $p = .14$). However, the interaction between condition and word-status was significant ($z = -3.88$, $p < .001$). Similar to Experiment 1, a baseline was created by averaging (fake) attended and unattended words in the control condition for contrast testing. This baseline was significantly higher than the error rate on the stressed words in the experimental condition ($z = 3.432$; $p < 0.001$), and significantly lower than that of the unstressed words ($z = -2.604$; $p = 0.009$).

In summary, when participants were instructed to emphasize a certain word in a sequence of words, they made fewer errors on that word, and more errors on the other words, compared to when no particular word was stressed. Is this pattern of results due to putting more emphasis on the word, or would any kind of manipulation to a word's production yield comparable results? In Experiment 3, we answer this question by asking the participants to do the reverse of emphasis; we asked them to silently mouth one of the four words.

Experiment 3

Methods

Participants: 39 undergraduate students at the University of Illinois at Urbana-Champaign, who had not taken part in experiments 1 or 2, participated in the study for course credit.

Materials: The same materials as before were used.

Procedure: The procedure was the same except for the instructions in the experimental condition. Participants were instructed to mouth the bold and underlined word, meaning that they had to make the gesture of saying the word with their mouth, without producing sounds.

Results and Discussion

The exclusion criteria were similar to Experiment 2, except that mouthing, instead of stressing, was the requirement by

which performance was judged. This procedure proved to be more difficult than the other two, and four participants could not follow the mouthing instructions even after going through the practice block twice. Moreover, there were many more silences in the control condition especially for participants who had first completed the mouthing condition, which led to the elimination of 7 trials (112 error opportunities) in this condition, over the remaining 35 participants. Silence or failure to properly mouth the word also led to the elimination of 271 error opportunities in the experimental condition.

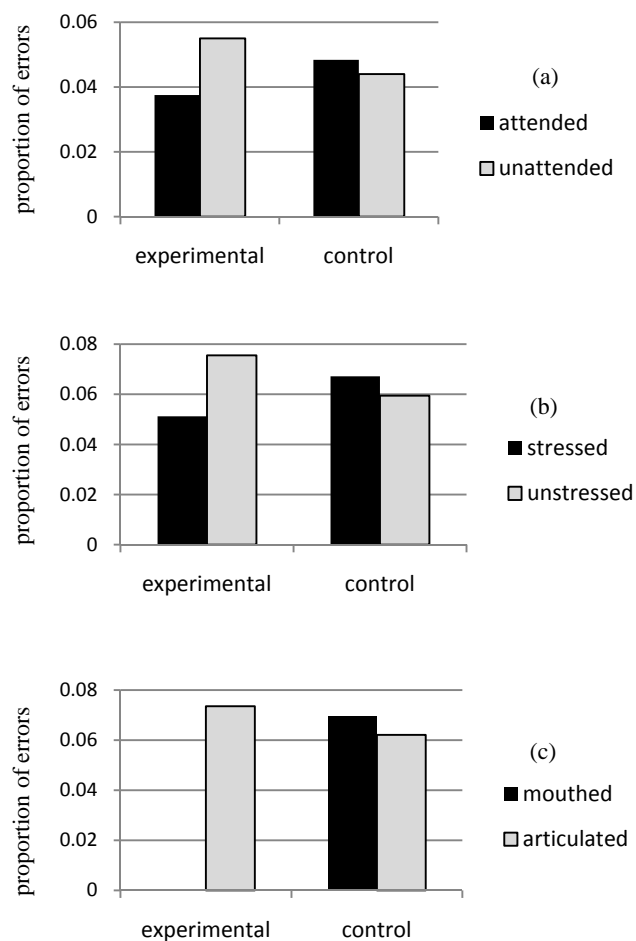


Figure 1(a-c) - Proportion of errors for each word-status (attended vs. unattended, stressed vs. unstressed, mouthed vs. articulated) in the experimental and control conditions in Experiments 1, 2 and 3.

Words were assigned to the fake mouthed status in the control condition in the fashion described for the other two experiments. Participants made 969 errors on the articulated words in the experimental condition and 832 in the control condition. Figure 1c shows the proportion of errors to the total error opportunities for each word-status in each condition. Note that there is no bar for the errors on the

mouthing words in the experimental condition, because by definition, these words were not supposed to be pronounced aloud (that recitation was excluded if they were), and therefore, there were no objective measures to decide whether an error was made or not. But the important comparison is between the proportion of errors on the articulated words in the experimental and control conditions, which was assessed using a hierarchical logistic model as before. Only the data for the articulated words were entered and a final model was built with condition as fixed effect, and random intercepts for subjects and items, as well as random slopes for the effect of condition on both subjects and items as random effects. This model showed that there were significantly more errors on the articulated words in the experimental than the control condition ($z = 2.19$; $p = .029$).

General Discussion

This paper posed the question of what happens if you focus attention on one word in a sequence of words. In three experiments, participants were asked to avoid mistakes, emphasize, or mouth one of the four words in tongue-twisters. Experiments 1 and 2 provided direct evidence that the experimental manipulation has been successful: there were fewer errors on the words that were singled out in some manner. However, this decrease in error rate on the attended/stressed words did not cause an overall lower error rate in the experimental condition. On the contrary, in both experiments, there were numerically more errors in that condition (although the difference did not reach statistical significance). This means that when participants focused their attention on one word, they made more errors on the other words in that sequence.

Intuitively, it seems natural for people to emphasize a word if they are more focused on it, so the purpose of experiment 3 was to test if the results could be replicated with attention implemented in the reverse way, namely by de-emphasizing articulation to an extreme extent, so that no sound is produced at all. If attention, rather than emphasis, is the factor that modulates the error rate, we expected to observe more errors on the articulated (unattended) words in the experimental, compared to the control, condition. The data supported this prediction. We conclude that regardless of the overt manifestation (e.g. making a word more or less prominent during articulation), focusing on a single word in a sequence hurts the production of other words.

In summary, attention matters. When a word is singled out in any respect, the production of other words suffers. An intuitive account is that greater attention means more activation on the attended word. In most production models, greater activation on the word promotes its accurate production, but increases the chance that that word will harm production by creating anticipatory and perseveratory errors on the other words (Dell, 1986). The constituents of the strongly-activated word will migrate in time to the unattended ones because of the relative disparity in activation. Alternately, the greater activation associated with

the attended word could mean less activation in absolute terms for the other words, assuming a limited pool of activation, or active inhibition of what is not attended.

Acknowledgments

This research was funded by a grant from the National Institutes of Health's National Institute for Deafness and Other Communication Disorders: DC000191.

References

- Allport, D. A. (1977). On knowing the meaning of words we are unable to report: The effects of visual masking. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, NJ: Lawrence Erlbaum.
- Davis, C. J., & Coltheart, M. (2002). Paying attention to reading errors in acquired dyslexia. *Trends in Cognitive Sciences*, 6, 359–361.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.
- Dell, G. S. & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, 20, 611–629.
- Friedmann, N., Kerbel, N., & Shvimer, L. (2010). Developmental attentional dyslexia. *Cortex*, 46(10), 1216–1237.
- Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, 47, 27–52.
- Humphreys, G. W., & Mayall, K. A. (2001). A peripheral reading deficit under conditions of diffuse visual attention. *Cognitive Neuropsychology*, 18, 551–576.
- Mayall, K. A., & Humphreys, G. W. (2002). Presentation and task effects on migration errors in attentional dyslexia. *Neuropsychologia*, 40, 1506–1515.
- McClelland, J. L., & Mozer, M. C. (1986). Perceptual interactions in two-word displays: Familiarity and similarity effects. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 18–35.
- Mozer, M. C. (1983). Letter migration in word perception. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 531–546.
- Nozari, N., Kittredge, A.K., Dell, G.S., Schwartz, M.F. (2010). Naming and repetition in aphasia: Steps, routes and frequency effects. *Journal of Memory and Language*, 63(4), 541–559.
- Oppenheim, G. M., & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition*, 106, 528–537.
- Saffran, E. M., & Coslett, B. (1996). “Attentional dyslexia” in Alzheimer’s disease: A case study. *Cognitive Neuropsychology*, 13, 205–228.
- Shallice, T., & McGill, J. (1978). The origins of mixed errors. In Requin J (Ed), *Attention and Performance VII*. Hillsdale, NJ: Erlbaum, 193–208.
- Shallice, T., & Warrington, E. K. (1977). The possible role of selective attention in acquired dyslexia. *Neuropsychologia*, 15, 31–41.

- Schwartz, M. F., & Hodgson, C. (2002). A new multiword naming deficit: Evidence and interpretation. *Cognitive Neuropsychology*, *19*(3), 263–288.
- Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97- 136.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*, 107-141.