Check for updates

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2016
Vol. 69, No. 9, 1719–1740, http://dx.doi.org/10.1080/17470218.2015.1093007

Routledge
Taylor & Francis Group

Executive control in spoken noun-phrase production: Contributions of updating, inhibiting, and shifting

Katarzyna Sikora[1], Ardi Roelofs[1], Daan Hermans[2,3], and Harry Knoors[2,3]

[1]Donders Institute for Brain, Cognition and Behaviour, Centre for Cognition, Radboud University, Nijmegen, The Netherlands
[2]Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands
[3]Royal Dutch Kentalis, Sint-Michielsgestel, The Netherlands

(Received 29 September 2014; accepted 2 September 2015; first published online 24 November 2015)

The present study examined how the updating, inhibiting, and shifting abilities underlying executive control influence spoken noun-phrase production. Previous studies provided evidence that updating and inhibiting, but not shifting, influence picture-naming response time (RT). However, little is known about the role of executive control in more complex forms of language production like generating phrases. We assessed noun-phrase production using picture description and a picture–word interference procedure. We measured picture description RT to assess length, distractor, and switch effects, which were assumed to reflect, respectively, the updating, inhibiting, and shifting abilities of adult participants. Moreover, for each participant we obtained scores on executive control tasks that measured verbal and nonverbal updating, nonverbal inhibiting, and nonverbal shifting. We found that both verbal and nonverbal updating scores correlated with the overall mean picture description RTs. Furthermore, the length effect in the RTs correlated with verbal but not nonverbal updating scores, while the distractor effect correlated with inhibiting scores. We did not find a correlation between the switch effect in the mean RTs and the shifting scores. However, the shifting scores correlated with the switch effect in the normal part of the underlying RT distribution. These results suggest that updating, inhibiting, and shifting each influence the speed of phrase production, thereby demonstrating a contribution of all three executive control abilities to language production.

Keywords: Executive control; Language production.

Individuals learn to produce language long before they acquire other cognitive skills, such as solving mathematical equations or playing chess. However, the mechanisms underlying fundamental language production abilities are not simple. To begin with, in producing an utterance, a speaker needs to select and encode appropriate words. A well-known theory of word production distinguishes between two stages: lexical selection and word-form encoding (Levelt, 1989, 2001; Levelt, Roelofs, & Meyer, 1999). During the lexical selection stage, a lemma representing the syntactic properties of the word is retrieved, while during the word-form encoding stage the appropriate morphophonological and phonetic form is encoded (Levelt, 2001). Lexical selection seems to

Correspondence should be addressed to Katarzyna Sikora, Donders Institute for Brain, Cognition and Behaviour, Centre for Cognition, Radboud University, Montessorilaan 3, Nijmegen 6525 HR, The Netherlands. E-mail: K.Sikora@donders.ru.nl

We thank Simon Fisher and two anonymous reviewers for helpful comments.

© 2015 The Experimental Psychology Society

be subject to competition in that a number of coactivated lemmas compete for selection, and their level of activation determines the speed of word production (Levelt, 2001; Piai, Roelofs, Jensen, Schoffelen, & Bonnefond, 2014; Roelofs, 1992, 2003). Moreover, in order to successfully complete an intended utterance, a speaker may need to allocate working memory capacity to the planning stages, monitor whether the internal or actually produced speech matches the intended utterance, and ignore environmental distractors (Roelofs, 2003, 2004; Roelofs & Piai, 2011). All these processes require some form of executive control, which refers to the higher level processes that regulate our perceptions, thoughts, and actions (e.g., Baddeley, 1996; Barkley, 2012; Gilbert & Burgess, 2008; Logan, 1985; D. E. Meyer & Kieras, 1997; for a recent overview, see Goldstein & Naglieri, 2014).

Executive control is an umbrella term covering a number of abilities. According to an influential proposal by Miyake and colleagues (Miyake et al., 2000), executive control consists of three main components: updating and monitoring of working memory representations, inhibiting of unwanted responses, and shifting between tasks or mental sets (see also Collette et al., 2005; Friedman et al., 2008). Updating is the ability to monitor incoming information for relevance to the task at hand, and maintain or actively manipulate the contents of working memory (e.g., Ecker, Lewandowsky, Oberauer, & Chee, 2010). Inhibiting is the ability to lower the activation of unwanted dominant, automatic, or prepotent responses (e.g., Aron, Robbins, & Poldrack, 2004). Shifting is the ability to switch back and forth between multiple tasks, operations, or mental sets (e.g., Allport & Wylie, 2000; Monsell, 2003). In the context of language production, updating may include the ability to manage communicative goals as well as monitor and schedule conceptual and linguistic processes (e.g., Levelt, 1989). For example, in a picture-naming task, the updating ability may determine how well a speaker keeps in mind the requirements of the task (e.g., to be fast and accurate) while engaging in conceptual and linguistic processes. The contribution of the inhibiting ability to language production may include the suppression of coactivated but incorrect lemmas. In order to correctly name a picture, a speaker may need to inhibit activated words different from the picture name (e.g., *dog*, *tail*, or *kitten* in naming a cat). The shifting ability may be engaged in language production, for example, when a speaker needs to switch from planning one type of phrase or sentence to another, from planning an utterance to monitoring output, or from planning an utterance to listening to a conversation partner.

Several types of evidence suggest that impaired language performance is associated with deficits in executive control, suggesting that executive control is important for skilled language performance. For example, individuals with developmental language disorder (Leonard et al., 2007; Montgomery, Magimairaj, & Finney, 2010), attention-deficit/hyperactivity disorder (ADHD; Bruce, Therlund, & Nettelbladt, 2006), Parkinson's disease (Altmann & Troche, 2011; Bastiaanse & Leenders, 2009), or stroke-induced brain damage (Brownsett et al., 2014; Schnur et al., 2009) experience language problems that may be related to problems in executive control, to varying extents and in different ways. Moreover, recent studies have demonstrated that individual differences in the speed of language production in healthy young speakers are also related to differences in executive control abilities.

Shao, Roelofs, and Meyer (2012) presented evidence for involvement of different executive control components in picture naming in healthy participants. They found that better updating and inhibiting abilities lead to shorter response times (RTs) for picture naming. More specifically, better updating ability was associated with shorter mean RTs for action naming, while better inhibiting ability was associated with shorter mean RTs for both object and action naming. Picture-naming RTs did not correlate with shifting ability. Analyses of the RT distributions suggested differential engagement of updating and inhibiting abilities across the experimental trials. RTs are not normally distributed but have a positive skew—that is, the tail is longer on the right than on the left side. The

longer right tail of the distribution reflects the proportion of the slowest responses, while the remaining responses are in the normal part of the distribution. Shao et al. observed that updating ability was correlated with the length of the right tail of the distribution of both action and object naming (i.e., the slowest responses), while inhibiting ability was correlated with the normal part of the RT distribution of action naming and the tail of the distribution of object naming.

These findings suggest that both updating and inhibiting play important roles in picture naming but in different ways. Updating ability is reflected in individual differences in picture-naming RT for the slowest responses, for both objects and actions. This result suggests that updating ability is especially taxed on the most demanding trials of picture naming. In contrast, inhibiting ability seems to be involved on most of the trials in action naming but only on the demanding trials in object naming. This may be due to the fact that action naming is generally more demanding than object naming, as reflected in longer mean RTs for the former. As mentioned, Shao et al. (2012) did not find a correlation between shifting ability and picture-naming RT. However, this does not exclude the possibility that shifting ability is involved in forms of language production that are more complex than picture naming, especially when actual switching is required. To conclude, Shao et al. found evidence that updating and inhibiting contribute to the speed of picture naming, in particular when naming is highly demanding, as reflected in the long RTs.

Piai and Roelofs (2013) assessed object-naming performance as part of a dual-task procedure and found evidence that the engagement of executive control depends on processing demand. In contrast to Shao et al. (2012), they did find a correlation between updating ability and mean object-naming RT. This result suggests that object naming taxes updating ability more in dual-task than in single-task performance, such that individual differences in this ability are reflected in the mean naming RTs.

To summarize, previous studies provided evidence for contributions of executive control to picture naming. To assess whether these findings generalize to more complex language performance, it is necessary to investigate whether each hypothesized component of executive control influences other forms of language production than picture naming. Moreover, it is important to assess whether shifting ability contributes to language production in situations that more explicitly engage this ability. The aim of the present study was to examine these issues by investigating the contributions of updating, inhibiting, and shifting to spoken phrase production in healthy adult participants.

Outline of the present study

We assessed phrase production performance using picture description and a picture-word interference procedure. First, we measured length, distractor, and switch effects in the picture description RTs, which were hypothesized to relate to updating, inhibiting, and shifting abilities, respectively. Second, for each participant we obtained scores on executive control tasks that measure verbal and nonverbal updating, nonverbal inhibiting, and nonverbal shifting abilities. We then tested for correlations between these executive ability scores and the magnitudes of the corresponding RT effects in the picture description task.

Participants described pictures of simple objects by producing noun phrases in Dutch. We used a set of only four pictures (i.e., a fork, a plate, a glass, and a bottle) to reduce variance due to differences in processing speed among pictures and words. The pictures were presented on the screen for only 250 ms. The brief presentation served to increase the demands on executive control. The small number of pictures and their brief presentation are somewhat uncommon features of our task meant to improve our assessment of the contributions of executive control. The pictures were presented in colour (e.g., a black line-drawing of a fork in green colour on a white computer screen) or without colour (e.g., a black line-drawing of a fork in white on a white computer screen).

In response to the coloured pictures, participants produced determiner–adjective–noun phrases, with

the adjective referring to the colour, henceforth the *long* phrases. In response to the white pictures, they produced determiner–noun phrases without the adjective, henceforth the *short* phrases. In one block of trials, the determiners were gender-marked *definite* articles, "de" (for nouns with common gender) or "het" (for nouns with neuter gender), and in another block of trials, the determiner was an *indefinite* article without gender-marking, "een". Thus, we examined not only the production of definite noun phrases, as has been done before (e.g., Roelofs, 2003, 2006; Schriefers, 1992, 1993; Schiller & Caramazza, 2003), but also the production of indefinite phrases, which has been neglected in previous research. In the determiner–adjective–noun phrases with the indefinite article, the adjective is inflected to mark the gender of the noun. For example, the adjective *groen* becomes "groen" for neuter nouns (e.g., "een groen bord", a green plate) and "groene" for common nouns (e.g., "een groene vork", a green fork). In contrast, with definite articles, the article but not the adjective is marked for grammatical gender (e.g., "het groene bord" and "de groene vork"). In planning determiner–adjective–noun phrases, speakers have to conceptually identify not only the pictured object but also its colour, and they have to retrieve from long-term memory a corresponding noun and colour adjective. Moreover, they have to retrieve the indefinite article and inflect the colour adjective (in the indefinite-article blocks) or retrieve the appropriate gender-marked article (in definite-article blocks). Following this, a syntactic structure has to be computed (Levelt, 1989) or retrieved from memory (Vosse & Kempen, 2000), and the determiner, adjective, and noun have to be serially ordered. In contrast, in planning determiner–noun phrases, speakers have to conceptually identify the pictured object but not the colour, retrieve a corresponding noun and definite or indefinite article, compute or retrieve a syntactic structure, and serially order the words. Conceptual preparation and syntactic encoding are followed by morphophonological and phonetic encoding and, finally, articulation (e.g., Levelt, 1989). Given that more information needs to be derived from the picture and manipulated in working memory for the long phrases than for the short phrases, we expected that updating ability would be more strongly engaged when producing the former. Consequently, the magnitude of the difference in RT between these phrase types, henceforth the *length effect*, is expected to reflect a speaker's updating ability (Korvorst, Roelofs, & Levelt, 2006; A. S. Meyer, Roelofs, & Levelt, 2003).

To assess the contribution of inhibiting ability, the pictures were combined with auditory distractor words, which could be congruent (e.g., the auditory distractor word "vork" combined with a picture of a fork) or incongruent (e.g., the auditory distractor word "bord" combined with a picture of a fork). We expected that inhibiting would be more strongly engaged with incongruent distractors than with congruent distractors and that the magnitude of the difference in RT between these distractor types, henceforth the *distractor effect*, would reflect a speaker's inhibiting ability. Note that the grammatical gender of the picture name and the incongruent distractor word may be the same or different, creating an additional source of interference, at either the level of the articles ("de" versus "het") or the inflection of the adjective (e.g., "groen" versus "groene"). The difference in RT between incongruent distractors with the same or different gender is referred to as the *gender congruency effect*.

As suggested by Schriefers, Meyer, and Levelt (1990) in a seminal article on the effect of spoken distractors in picture naming, there exists no good neutral baseline. For example, presenting a spoken word distractor will delay picture naming relative to a silence condition regardless of the relationship between picture and distractor simply because of the sound of the spoken distractor. However, presenting pink noise instead of using silence will suffer from the problem of comparing the effect of linguistic and nonlinguistic distractors. Therefore, the psycholinguistic literature has settled on testing between different word conditions, such as semantically related versus unrelated, phonologically related versus unrelated, or incongruent (semantically related) versus congruent (identical), as is often done in the literature on

Stroop-like effects (e.g., colour–word Stroop, Eriksen flanker, or Simon, see Piai, Roelofs, Acheson, & Takashima, 2013). We decided to use incongruent and congruent distractors, because this contrast typically yields the largest and most robust RT effect (e.g., Schriefers et al., 1990), which would increase the chance of finding a correlation with inhibiting ability. The incongruent–congruent contrast does not allow us to specifically relate the effect to any stage of word production (unlike semantic or phonological effects), but this was also not the aim of our study. Table 1 gives example materials for the length and distractor conditions for each of the grammatical genders and definiteness conditions.

To assess the contribution of the shifting ability, the required phrase type (long or short) changed every second trial. Thus, two short phrases (for pictures in white) could be followed by two long phrases (for pictures in colour) or vice versa. A trial that repeats the previous phrase type (short preceded by short or long preceded by long) is referred to as a repeat trial, and a trial that does not repeat the previous phrase type (short preceded by long or long preceded by short) is referred to as a switch trial. Speakers may need to engage the shifting ability on switch trials to enable to production of a different phrase type. We expected that shifting would be more strongly engaged on switch than on repeat trials and that the magnitude of the difference in RT between these trial types, henceforth the *switch effect*, would reflect a speaker's shifting ability.

In addition, we measured participants' updating, inhibiting, and shifting abilities with standard tasks commonly used to assess executive control (Collette et al., 2005; Conway et al., 2005; Friedman et al., 2008; Miyake et al., 2000). It could be possible that there are linguistic updating processes that are partly independent from the nonverbal processes (e.g., Montgomery et al., 2010). Therefore, we used the operation-span and odd-one-out tasks (Conway et al., 2005) to assess verbal and nonverbal updating ability, respectively. Moreover, we used the stop-signal task (Verbruggen, Logan, & Stevens, 2008) to assess nonverbal inhibiting ability, and the shape–colour switching task (Miyake et al., 2000) to assess nonverbal shifting ability. Detailed descriptions of these tasks are given below in the Method section. We expected that individual differences in the scores for these executive control tasks would correlate with the corresponding effects in the picture description RTs. In particular, the operation-span and odd-one-out scores were expected to correlate with the length effect in RTs (both reflecting the updating ability), the stop-signal scores were expected to correlate with the distractor effect (both reflecting the inhibiting ability), and the scores on the shape–colour switching task were expected to correlate with the RT switch effect (both reflecting the shifting ability).

Following Shao et al. (2012), we examined correlations of executive control scores not only with mean RTs but also with components of the underlying RT distribution by performing ex-Gaussian analyses. The ex-Gaussian function consists of a convolution of a Gaussian (i.e., normal) and an exponential distribution, which generally provides good fits to empirical RT distributions (e.g., Luce, 1986; Ratcliff, 1979). The analyses provide three parameters characterizing a distribution, called μ, σ, and τ. The μ and σ parameters reflect the mean and standard deviation of the Gaussian portion, and τ reflects the mean and standard deviation of the exponential portion. The mean of the whole distribution (i.e., the mean RT) equals the sum of μ and τ. Thus, ex-Gaussian analyses decompose mean RTs into two additive components, which characterize the normal part (μ) and the right tail (τ) of the underlying RT distribution.

Finally, following Shao, Meyer, and Roelofs (2013), we performed delta-plot analyses to further examine the contribution of inhibiting. Previous research distinguished between nonselective and selective inhibition (Forstmann et al., 2008). Nonselective inhibition is involved in the suppression of any unwanted response (like in the stop-signal task or any distractor in the picture–word interference task), while selective inhibition is involved in the suppression of specific responses that compete for selection (like semantically related distractors in the picture–word interference task). A way to measure selective inhibition is by calculation of delta plots. Delta plots can be

Table 1. *Example materials for the length and distractor conditions for each of the two grammatical genders and definiteness conditions*

		Length		*Distractor*		
Definiteness	*Gender*	*Short*	*Long*	*Congruent/ congruent*	*Incongruent/ congruent*	*Incongruent/ incongruent*
Definite article	Common	de vork (the fork)	de groene vork (the green fork)	vork (fork)	fles (bottle)	bord (plate)
	Neuter	het bord (the plate)	het groene bord (the green plate)	bord (plate)	glas (glass)	vork (fork)
Indefinite article	Common	een vork (a fork)	een groene vork (a green fork)	vork (fork)	fles (bottle)	bord (plate)
	Neuter	een bord (a plate)	een groen bord (a green plate)	bord (plate)	glas (glass)	vork (fork)

Note: Length: short, long; distractor: congruent noun and gender, incongruent noun with congruent gender, incongruent noun with incongruent gender; grammatical gender: common, neuter; definiteness condition: definite, indefinite. English translations of the Dutch materials are shown in parentheses.

derived by rank ordering condition RTs and dividing them into quantiles (e.g., 20% bins). The distractor effect or "delta" is then determined for each quantile—that is, as a function of the relative speed of responding. Evidence suggests that selective inhibition builds up over time, so that the influence of inhibition is most strongly reflected in the slowest responses (e.g., Ridderinkhof, 2002; Van den Wildenberg et al., 2010). That is, better inhibitory ability is associated with smaller distractor effects for the slowest responses. This may be quantified by computing the slopes of the delta plots—that is, the difference between the deltas of consecutive quantiles divided by the difference in the corresponding quantile mean RTs. Previous research has suggested that better inhibiting ability is associated with shallower or negative-going slopes for the slowest responses—that is, the slopes of the slowest delta segments (e.g., Ridderinkhof, 2002; Van den Wildenberg et al., 2010).

Shao et al. (2013) obtained evidence that both selective and nonselective inhibition are involved in picture naming but in different ways. Participants had to name pictured objects (e.g., a fork) with superimposed written distractor words, which were semantically related (e.g., *knife*) or unrelated (e.g., *car*). Nonselective inhibiting ability, as indexed by scores for the stop-signal task, was correlated with the mean RT regardless of the type of word distractor (i.e., semantically related or unrelated) in picture naming. In contrast, selective inhibiting ability, as indexed by slopes of the slowest delta segment, was correlated with the magnitude of the semantic effect (i.e., the difference in RT between semantically related and unrelated distractors). Moreover, the selective and nonselective inhibiting measures were not correlated. This suggests that selective inhibiting is engaged to selectively suppress distractors that are semantically related (thereby reducing the magnitude of the semantic effect), whereas nonselective inhibiting is engaged to suppress any competing word, regardless of semantic relatedness (thereby reducing the mean RTs). In the present study, we used the stop-signal task to assess nonselective inhibiting ability, and we performed delta-plot analyses to assess selective inhibiting ability.

EXPERIMENTAL STUDY

Method

Participants

Forty native speakers of Dutch participated in the experiment (28 women and 12 men, mean age =

24.08 years, age range: 19 to 37 years). The participants were recruited via the Radboud University SONA system. They received 12.50 euros or 1.5 credits for their participation.

Procedure and design

The participants signed the informed consent before the experimental session. Following standard practice in individual differences research, tasks and blocks within tasks were presented to all the participants in the exact same order to minimize any measurement error due to participant by order interaction (cf. Friedman et al., 2008; Miyake et al., 2000). First, they performed the picture description task, followed by the odd-one-out task, the stop-signal task, the operation-span task, and the shape–colour switching task. After the experimental session, participants were debriefed and paid for their participation. An experimental session lasted about 1 hour and 15 minutes.

Picture-description task. In this task, participants had to describe a picture presented in the middle of a computer screen while trying to ignore a spoken distractor that was played via headphones. Each trial began with a fixation cross, which remained on the screen for 700 ms, followed by the presentation of a spoken distractor and a picture simultaneously (i.e., with the same presentation onset). The picture remained on the screen for 250 ms followed by a blank screen for 2150 ms. Each picture was presented with a congruent distractor, an incongruent distractor with congruent gender, and an incongruent distractor with incongruent gender. Hence, each picture was pared with three different spoken distractors. Each spoken distractor was presented an equal number of times in the experiment. For each participant, the stimulus list was randomized using the programme Mix (Van Casteren & Davis, 2006) with the restriction that stimuli were not repeated on consecutive trials. To focus on the influences of executive control and to reduce variance due to differences in processing speed among pictures and words, only a limited number of objects and colours were used. The set of stimuli consisted of four pictures—namely, a bottle, a plate, a glass, and a fork, and four spoken distractors, which were the names of these objects. All spoken distractors were monosyllabic Dutch words: *fles* (*bottle*), *bord* (*plate*), *glas* (*glass*), and *vork* (*fork*). The words were all semantically related (like in the colour–word Stroop task, e.g., Roelofs, 2003). The spoken words were recorded by a female native speaker of Dutch. The mean duration of the spoken distractor was 530 ms. All the participants ran through the experiment in the exact same order (cf. Friedman et al., 2008; Miyake et al., 2000). There were two practice blocks and eight experimental blocks of trials. Each experimental session began with two practice blocks. In the first practice block, participants had to describe all the pictures using noun phrases with a definite article, and in the second practice block participants described all the pictures using phrases with an indefinite article. Next, in the experimental part, participants used definite phrases (definite condition) in four of the blocks and indefinite phrases (indefinite condition) in another four blocks. Participants always began with a definite article block followed by an indefinite article block. The four definite and four indefinite blocks of trials were alternated: definite, indefinite, definite, indefinite, and so on. Before each experimental block, participant received information about the required type of noun phrase, definite or indefinite. Each experimental block consisted of 48 trials. In total there were 192 trials for the definite condition and 192 trials for the indefinite condition.

As mentioned earlier, the Dutch language distinguishes between common nouns, which take the article "de", and neuter nouns, which take the article "het" as definite determiners. Two words, *glas* (*glass*) and *bord* (*plate*), were neuter-gender nouns, and the two other words, *vork* (*fork*) and *fles* (*bottle*), were common-gender nouns. The picture names and the spoken distractors could be two identical nouns (congruent condition), two different nouns taking the same grammatical gender (incongruent noun with congruent gender condition), or two different nouns taking different genders (incongruent noun with incongruent gender condition). Each of these distractor

conditions had the same number of trials. The pictures were either black-and-white drawings or coloured drawings, either blue or green. The participants were instructed to produce determiner–noun phrases (e.g., "de vork" or "een vork") when the presented picture was a black-and-white drawing (the short phrase condition). When the picture was presented in one of the two colours, the participants had to produce a phrase that included an article, a colour adjective, and the name of the object (e.g., "de groene vork" or "een groene vork", the long phrase condition). The number of trials for these conditions was the same. The pictures were presented such that the required phrase type changed every second trial. Thus, in the trial sequence, two black-and-white pictures were followed by two coloured pictures, which were followed by two black-and-white pictures, and so on. This designed allowed us to measure picture description RT on repeat and switch trials. The number of trials for the repeat and switch conditions was equal.

Odd-one-out task. The odd-one-out task measures nonverbal updating ability (Conway et al., 2005). The odd-one-out task consisted of 60 triples of drawings representing arbitrary shapes. For each triple, two shapes were identical, and one was different. The three figures were presented on the computer screen, and the participants were instructed to indicate by pressing one of three buttons which figure is different from the others (i.e., the odd-one-out). The left button denoted the left figure, the middle button denoted the middle figure, and the right button denoted the right figure. Moreover, the participants were told to remember the location of the odd-one-out figure on each trial. After a random number of trials (varying between two and six), the participants had to recall the location of all odd-one-out figures since the beginning of a set. A table with three columns was presented on the computer screen. Each column referred to the three locations, and the rows referred to the trials. The participants indicated the recalled locations of the odd-one-out figures in the table, in the correct order, by sequentially pressing the corresponding buttons. This was followed by a next set of new figures.

Operation-span task. The operation-span task measures verbal updating ability (Conway et al., 2005). The operation-span task consisted of 60 mathematical operations and 60 Dutch words. Materials were presented on the computer screen. Each trial began with a fixation cross presented for 800 ms followed by a mathematical operation and a word presented in the middle of the screen (e.g., 12 − 4/2 = 3? Strand). The participants were instructed to read aloud the mathematical operation and the word. Next, the participants had to indicate by a button press whether the mathematical operation was correct. After a random number of trials (varying between two and six), participants had to recall all words presented since the beginning of a set. The instruction "write down" displayed on the screen indicated that a participant had to recall and write down all words in the correct order. This was followed by a new set of mathematical operations and words.

Stop-signal task. The stop-signal task measures nonverbal inhibiting ability (Verbruggen et al., 2008). The task consisted of 75% "go" trials and 25% "stop" trials. Each go trial began with a fixation point presented in the middle of the screen for 250 ms, followed by a target stimulus. The target stimulus was either a square or a circle. The participants were instructed to respond to the stimuli by pressing one button (labelled "?") when they saw a circle and another button (labelled "Z") when they saw a square. The stimuli remained on the screen until the participant responded but not longer than for 1250 ms. The participants were told to respond to the stimuli as quickly and accurately as possible. On the stop trials also an auditory stimulus was presented. The auditory stimuli followed the visual stimuli. The participants were instructed to inhibit their response to the visual stimuli on the trials when the auditory stimuli were presented. First the auditory stimuli were presented 250 ms after onset of the visual stimuli (the stop-signal delay). After each successful stop trial, the stop-signal delay was increased

with 50 ms, while after each unsuccessful stop trial the stop-signal delay was decreased with 50 ms. The task consisted of one practice and three experimental blocks. The practice block included 32 trials, and each experimental block included 64 trials.

Shape–colour switching task. The shape–colour switching task measures nonverbal shifting ability (Miyake et al., 2000). The task consisted of visual stimuli presented on the computer screen. The stimuli were a square or a circle that could be either red or green. The participants were instructed to respond to the colour of the stimuli when the figures were presented in the top of the screen, and to the shape of the stimuli when the figures were presented in the bottom of the screen. Response buttons were "↓" and "↑". The participants had to press "↓" as a response to the stimuli that were either red or a circle, and they had to press "↑" as a response to the stimuli that were either green or a square. There were three practice blocks and three experimental blocks. In the first block, stimuli were presented only in the top of the screen, the colour block. In the second block, stimuli were presented only in the bottom of the screen, the shape block. In the last block, stimuli were presented in clockwise rotation, beginning in the upper left quadrant of the screen, followed by the upper right quadrant, followed by the lower right quadrant, and the lower left quadrant. The shape and the colour blocks consisted of 48 trials. The mixed block consisted of 128 trials.

Data analysis

Picture-description task. The data of four participants for this task were excluded from the analysis because they did not correctly follow the instructions. In total, the data of 36 subjects were included in the analysis. Responses were excluded from the analysis if the produced phrase did not match the correct phrase or when the response included any kind of disfluency or was not completed before the end of a trial. Mean RTs were calculated separately for the definite blocks and indefinite blocks. For each block, mean RTs were calculated for seven conditions: long phrase condition, short phrase condition, congruent condition, incongruent noun with congruent gender condition, incongruent noun with incongruent gender condition, repeat condition, and switch condition. Repeated measures analyses of variance (ANOVAs) were conducted to test for main effects and interactions between conditions. Four effects were defined: *definiteness* (definite vs. indefinite), *length* (short vs. long phrase), *distractor* (congruent vs. incongruent noun with congruent gender vs. incongruent noun with incongruent gender), and *switch* (repeat vs. switch trials).

Pearson correlations were computed between the scores for the executive control tasks and the three effects measured by the language production task (length, distractor, and switch). The length effect was calculated for each participant by subtracting the mean RT of the short phrase condition from the mean RT of the long phrase condition. The distractor effect was calculated by subtracting the mean RT of the congruent condition from the mean RT of the incongruent noun with congruent gender condition. Note that the incongruent noun with incongruent gender condition was not entered into the correlation analyses. Whereas the congruent condition and the incongruent noun with congruent gender condition differ in the congruency of the noun but not the grammatical gender (which is the same for picture name and distractor word), the grammatical gender differs between the congruent condition and the incongruent noun with incongruent gender condition. Thus, the latter contrast includes a difference not only in noun but also in grammatical gender, which complicates the interpretation of the effect. Therefore, the distractor effect that entered the correlation analyses concerned the difference in RT between the congruent condition and the incongruent noun with congruent gender condition. Finally, the switch effect was calculated by subtracting the mean RT of the repeat condition from the switch condition.

Odd-one-out task. Scores for the odd-one-out task were calculated following the guidelines of Conway et al. (2005). The score of one participant was excluded from the analysis because the accuracy

in the odd-one-out task was below 85%. The number of correctly recalled locations of the odd figures in a set was calculated. The number of figures in a set varied between two and six. The participants received 1 point for each correctly recalled set. There were in total 15 sets. Thus, the total scores could range between 0 and 15. The score for each set was calculated as the proportion of the correctly recalled locations and the total number of locations to be recalled within the set. For example, correctly recalling the locations of two odd figures of a set of six figures was scored as 0.33 point, while correctly recalling the locations of two odd figures of a set of four figures was scored as 0.5 point. Higher total scores on the odd-one-out task indicate better (nonverbal) updating ability.

Operation-span task. Scores for the operation-span task were also calculated following the guidelines of Conway et al. (2005). The scores of two participants were excluded from the analysis as one of the participants had lower than 85% accuracy for the mathematical operations, and another participant did not correctly follow the instructions. The number of correctly recalled words for each set was calculated. The number of the words in each set varied between two and six. In total there were 15 sets. Participants could receive 1 point for each correctly recalled set; hence the total score could range between 0 and 15. The score for each set was calculated as the proportion of the correctly recalled words and the total number of words to be recalled within the set, similar to the scoring of the odd-one-out task. Higher total scores for the operation-span task indicate better updating ability.

Stop-signal task. Scores were calculated following the instructions of Verbruggen et al. (2008). The data of three participants were excluded from the analysis because of poor performance. The stop-signal reaction time (SSRT) was calculated for each participant. The SSRT is equal to the difference between the mean RT of all go-trials and the mean stop-signal delay. Shorter SSRTs indicate better inhibiting ability.

Shape–colour switching task. Trials were categorized as errors and excluded from the analysis if a participant failed to respond correctly to the presented figure. Mean RTs for the switch trials from the third experimental block (trials from the upper left and the lower right quadrants of the screen) and for the repeat trials from the third experimental block (trials from the upper right and lower left quadrants of the screen) were calculated. The shifting score was obtained by subtracting the mean RT for the repeat trials from mean RT for the switch trials. Lower values indicate better shifting ability.

Correlation analyses. Pearson correlations were computed between the scores of the executive control tasks and the length, distractor, and switch effects in the picture description RTs as well as the overall mean RTs. Moreover, correlations were computed between the executive scores and the length, distractor, and switch effects in the three ex-Gaussian parameters (μ, σ, and τ) as well as the mean values of the parameters across all conditions.

Ex-Gaussian analyses. For each participant, we estimated the values of the three ex-Gaussian parameters μ, σ, and τ for the RTs in each condition. The parameter values were estimated using the quantile method implemented in the QMPE software (Brown & Heathcote, 2003). We then computed the effects of length, distractor, and switch in these parameters. As explained earlier, the effects in the three ex-Gaussian parameters as well as the mean values of the parameters across all conditions were entered into the correlation analyses.

Delta-plot analyses. For each participant, the RTs of the congruent condition and the incongruent noun with congruent gender condition were rank-ordered and divided into quintiles (i.e., 20% bins), for each condition separately. For each quintile, mean RTs were calculated, and the interference effect (delta) for each quintile was computed. Finally, the slopes of the lines that connect the delta values of consecutive quintiles were computed

(Ridderinkhof, 2002). The literature suggests that the slope of the slowest delta segment (i.e., the slope of the line connecting the fourth and fifth quintiles) best indexes inhibition ability (e.g., Forstmann et al., 2008; Van den Wildenberg et al., 2010).

Results

Mean picture-description performance

The mean RTs and error rates for each of the conditions of the picture-description task are presented in Table 2. Mean RTs were longer for the definite than for the indefinite phrases (definiteness effect), shorter for the short than for the long phrases (length effect), shorter for the congruent condition than for the incongruent noun with congruent gender condition (distractor noun effect), shorter for the incongruent noun with congruent gender condition than for the incongruent noun with incongruent gender condition (gender congruency effect), and longer for switch than for repeat trials (switch effect).

In tests of sphericity of the RT data, Mauchly's test of sphericity was significant for the interaction of definiteness and distractor, $\chi^2(2) = 6.52$, $p < .05$, which indicates a violation of the assumption of sphericity. Therefore, Greenhouse–Geisser estimates of sphericity were used to correct the degrees of freedom ($\varepsilon = .85$).

All main RT effects were found to be significant at $p < .001$. There was a main effect of definiteness: The participants responded faster in the indefinite than in the definite condition, $F(1, 35) = 145.39$, $MSE = 22{,}801$, $\eta_p^2 = .81$. In addition, there was a main distractor effect, $F(2, 70) = 44.25$, $MSE = 9060$, $\eta_p^2 = .56$: Participants responded faster in the congruent noun condition than in the incongruent noun with congruent gender condition (distractor noun effect), $F(1, 35) = 44.85$, $MSE = 2085$, $\eta_p^2 = .56$, and slower in the incongruent noun with incongruent gender condition than in the incongruent noun with congruent gender condition (gender congruency effect), $F(1, 35) = 10.30$, $MSE = 1647$, $\eta_p^2 = .23$. There was also a main length effect: Participants responded faster in the short than in the long phrase condition, $F(1, 35) = 29.81$, $MSE = 30{,}827$, $\eta_p^2 = .46$. Finally, there was a main switch effect: The participants responded faster in the repeat than in the switch condition, $F(1, 35) = 24.57$, $MSE = 5682$, $\eta_p^2 = .41$.

Additionally, we found three significant interactions. There was an interaction between definiteness and distractor, $F(1.7, 59.6) = 11.8$, $MSE = 70{,}160$, $p < .001$, $\eta_p^2 = .25$. As Figure 1 shows, this interaction occurred because the incongruent distractors yielded a difference in RT between the

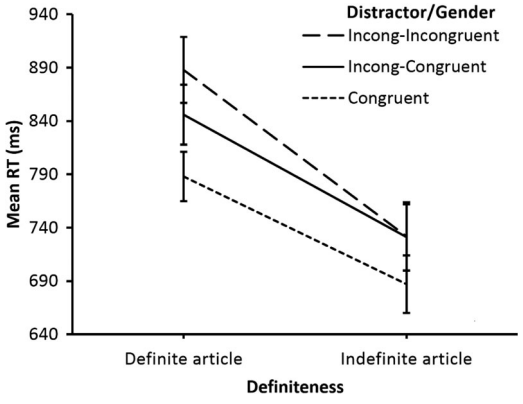

Figure 1. *Mean response time (RT) for the picture descriptions per distractor condition (congruent noun/congruent gender, incongruent noun/incongruent gender, incongruent noun/congruent gender) and definiteness condition (definite, indefinite). The error bars indicate one standard error.*

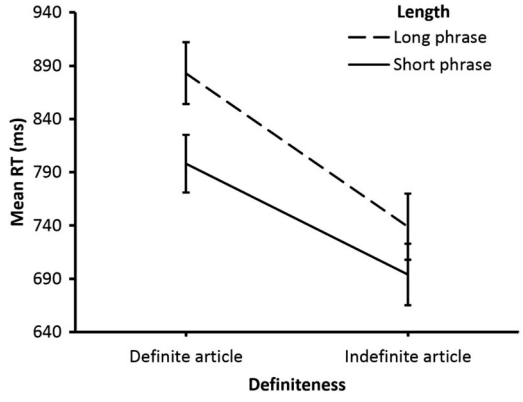

Figure 2. *Mean response time (RT) for the picture descriptions per phrase type (long, short) and definiteness condition (definite, indefinite). The error bars indicate one standard error.*

Table 2 *Mean reaction times and percentage errors in the picture-description task*

			Switch		
Definiteness condition	*Distractor/Gender*	*Length*	*Repeat*	*Switch*	*Total*
Definite article	Congruent/congruent	Long	845 (12)	819 (10)	831 (11)
		Short	714 (5)	773 (3)	744 (4)
		Total	779 (8)	796 (7)	788 (7)
	Incongruent/congruent	Long	891 (15)	877 (14)	884 (14)
		Short	778 (6)	838 (5)	808 (5)
		Total	834 (10)	857 (10)	846 (10)
	Incongruent/incongruent	Long	935 (21)	934 (19)	934 (20)
		Short	808 (12)	876 (9)	842 (10)
		Total	871 (17)	905 (14)	888 (15)
Total		Long	890 (16)	876 (14)	883 (15)
		Short	767 (6)	829 (7)	798 (6)
		Total	828 (11)	853 (10)	840 (10)
Indefinite article	Congruent/congruent	Long	728 (10)	701 (8)	714 (9)
		Short	629 (2)	690 (3)	660 (2)
		Total	678 (6)	696 (5)	687 (6)
	Incongruent/congruent	Long	739 (14)	760 (13)	750 (13)
		Short	677 (4)	746 (5)	712 (4)
		Total	708 (9)	753 (9)	731 (9)
	Incongruent/incongruent	Long	768 (14)	740 (14)	754 (14)
		Short	680 (5)	741 (8)	710 (6)
		Total	724 (9)	740 (11)	732 (10)
Total		Long	745 (13)	734 (11)	739 (12)
		Short	662 (3)	726 (5)	694 (4)
		Total	703 (8)	730 (8)	717 (8)
Overall	Congruent/congruent	Long	786 (11)	760 (9)	773 (10)
		Short	672 (3)	732 (3)	702 (3)
		Total	729 (7)	746 (6)	737 (7)
	Incongruent/congruent	Long	815 (14)	819 (13)	817 (14)
		Short	728 (4)	792 (5)	760 (5)
		Total	771 (9)	805 (9)	788 (9)
	Incongruent/incongruent	Long	851 (18)	837 (17)	844 (17)
		Short	744 (8)	809 (8)	776 (8)
		Total	797 (13)	823 (13)	810 (13)
Total		Long	817 (14)	805 (13)	811 (14)
		Short	714 (5)	777 (5)	746 (5)
		Total	766 (10)	791 (9)	779 (10)

Note: Response times in milliseconds. Error percentages in parentheses.

incongruent- and congruent-gender conditions for the definite phrases but not for the indefinite phrases. Moreover, there was an interaction between definiteness and length, $F(1, 35) = 23.94$, $MSE = 86{,}108$, $p < .001$, $\eta_p^2 = .41$. Figure 2 shows that this interaction occurred because the length effect was larger for the definite than for the indefinite phrases. Finally, there was an interaction between length and switch, $F(1, 35) = 33.9$, $MSE = 304{,}638$, $p < .001$, $\eta_p^2 = .49$. Figure 3 shows that this interaction occurred because the switch effect was present for the short phrases, $t = 8.79$, $p < .001$, but not for the long phrases, $t = -1.192$, $p = .24$.

Correlation analyses

We analysed the correlations among the scores of the executive control tasks (i.e., operation span, odd-one-out, stop signal, and shape–colour switching), the mean picture description RTs, and the three effects in the RTs (i.e., length, distractor, and switch). Excluded data points were coded as missing values. Moreover, the score of one participant on the operation-span task was identified as an outlier using the Mahalanobis distance (Conway et al., 2005) and was excluded from the analysis. The correlations are presented in Tables 3 and 4.

Table 3 presents the correlations among the scores on the executive control tasks. The table shows that there were no significant correlations except between the scores for the operation-span and odd-one-out tasks. Table 4 presents the correlations among the scores for the executive control tasks, the mean picture description RTs, and the three effects in the RTs. The table shows that the operation-span and odd-one-out scores correlated with the overall mean RT, and, additionally, the operation-span scores correlated with the length effect in the RTs. Moreover, the stop-signal scores correlated with the distractor effect. For the shape–colour switching scores no significant correlations were obtained.

In the correlation analyses, we did not include an equal number of participants for each test. Instead, we included the score of a participant if it met the criterion of a test (i.e., operation span task, odd-one-out task, stop-signal task, and colour–shape task) regardless of whether the participant met the criteria of all the other tests. For instance, if a participant had to be excluded from the stop-signal task because the score did not qualify according to the criteria of this test (e.g., too many errors), the scores of this participant on other tests were still included if these scores passed the criteria of those tests. If we exclude a participant's scores on all of the tests because the scores of one test did not pass the criteria of that test, the number of participants entering the correlation analyses would considerably decrease. This would likely result in weaker correlations. Still, to test whether the reported results were due to the unequal number of participants on the tests, we ran our analysis with an equal number of participants on all tests (i.e., using the general exclusion criterion) and thus a lower total number of participants. These correlations are presented between parentheses in Table 4. As can be seen, the correlations were overall weaker but still significant. Thus, the pattern of correlations is the same with unequal and equal numbers of participants for the executive control tasks.

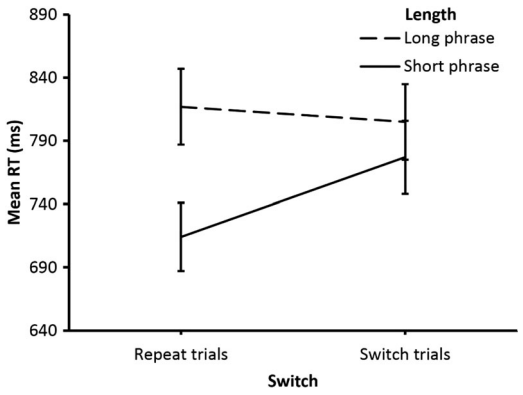

Figure 3. *Mean response time (RT) for the picture descriptions per length condition (short, long) on repeat and switch trials. The error bars indicate one standard error.*

Ex-Gaussian analyses

We also analysed the correlations between the scores for the four executive control tasks and the values of the three ex-Gaussian parameters across conditions. Moreover, we analysed the correlations between the executive control scores and the length, distractor, and switch effects in the three ex-Gaussian parameters characterizing the RT distributions.

We found correlations between the scores on the operation-span and odd-one-out tasks and the μ of the distribution of RTs across all conditions, $r = -.407$, $p < .01$, and $r = -.509$, $p < .001$, respectively. Moreover, we found a positive correlation between the scores on the shape–colour switching task and the switch effect in μ, $r = .316$, $p < .05$, while the correlation was negative for the switch effect in the τ parameter, $r = -.36$, $p < .05$. The negative correlation between the scores on the shape–colour switching task and the τ parameter was surprising. We address this finding in the Discussion section.

Table 3. *Correlations among the scores for the executive control tasks*

Task	*Odd-one-out*	*Operation span*	*Stop signal*
Operation span	.582**	—	—
Stop signal	−.212	−.014	—
Shape–colour switching	−.217	−.098	.085

Note: Odd-one-out and operation span are assumed to index updating ability, stop signal is assumed to index inhibiting ability, and shape–colour switching is assumed to index shifting ability.
**$p < .01$.

Delta-plot analyses

We analysed the correlations among the slope of slowest delta segment, the scores on the stop-signal task, and the mean distractor effect in the RTs.

The stop-signal scores correlated with the mean distractor effect in the RTs, $r = .451$, $p < .01$. Moreover, we found a correlation between the slopes of the slowest delta segment (i.e., the final slope) and the distractor effect, $r = .294$, $p < .01$. Figure 4 shows the corresponding scatter plots. There was no correlation between the stop-signal scores and the slopes of the slowest delta segment, $r = .073$, $p = .34$.

Discussion

The present study examined the influence of executive control on language production. More specifically, we investigated how individual differences in updating, inhibiting, and shifting abilities influence phrase production. The language task used in this study allowed us to measure the involvement of all three components of executive control proposed by Miyake et al. (2000). We measured how fast a speaker produces a phrase in higher demand updating, inhibiting, and shifting conditions as compared to lower demand conditions. We argued that executive control will be more strongly engaged in more demanding conditions, and hence differences in the speed of performance between higher and lower demand conditions will reflect differential engagement of the speakers' updating, inhibiting, and shifting abilities. We found that participants responded faster in the lower than in the higher demand conditions, as reflected in the mean RTs. More specifically, speakers described pictures faster in the short- than in the long-phrase condition, faster in the congruent than in the incongruent distractor conditions, and faster in the repeat than in the switch condition. These findings suggest that an increase in executive demand has consequences for the speed of language production.

On the long phrase trials, speakers had to process and maintain in working memory the relevant information for three components of the phrase—namely, the determiner, adjective, and noun, whereas on short phrase trials, this was required only for two components—namely, the determiner and noun. This difference in processing demand was reflected in the RTs as the length effect. Moreover, individual differences in verbal updating scores, obtained by the operation-span task, correlated with the magnitude of the length effect in the RTs. Better updating scores were associated with smaller length effects. Both verbal

Table 4. *Correlations among the scores for the executive control tasks and the mean reaction times and effects in the picture-description task*

Task	*Length effect*	*Distractor effect*	*Switch effect*	*Total RT*
Odd-one-out	.085 (−.033)	−.276 (−.214)	.046 (−.049)	−.510** (−.418*)
Operation span	−.353* (−.336*)	−.089 (−.092)	.057 (.047)	−.307* (−.338*)
Stop signal	−.079 (−.141)	.451** (.402*)	.121 (.153)	.186 (.167)
Shape–colour switching	.112 (−.142)	−.031 (.056)	.031 (.021)	.148 (.067)

Note: RT = response time. The correlations for participants with scores on all executive control tasks are shown in parentheses.
*$p < .05$. **$p < .01$.

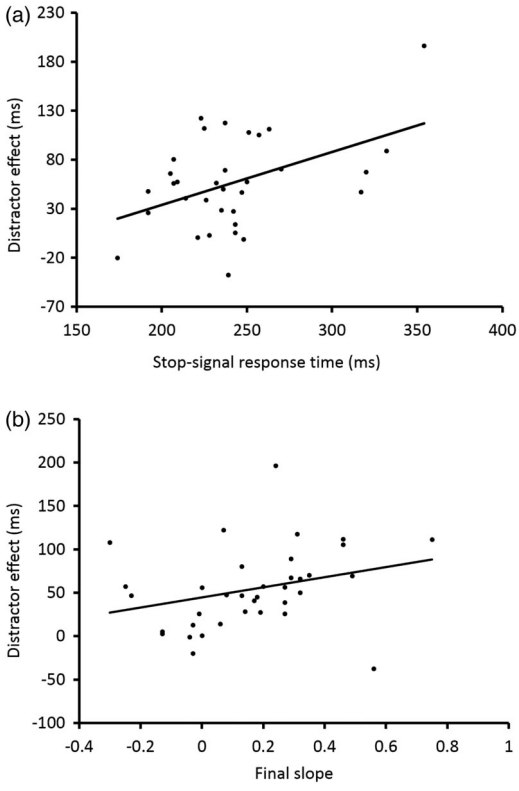

Figure 4. *Scatter plots of the relationship between (a) the distractor effect and the stop-signal response time and (b) the distractor effect and the slope of the slowest delta segment.*

and nonverbal updating scores (the latter obtained by the odd-one-out task) correlated with overall mean RTs. Better updating scores were associated with shorter RTs. These results suggest that the speed of phrase production is influenced by the updating ability of the speaker.

On incongruent distractor trials, speakers had to overcome interference of incongruent distractor words, whereas on congruent distractor trials, there was no such interference. This difference in processing demand was reflected in the RTs as the distractor effect. Moreover, individual differences in inhibiting scores, as obtained by the stop-signal task, correlated with the magnitude of the distractor effect in the RTs. Better inhibiting scores were associated with smaller distractor effects. These results suggest that phrase production is influenced by inhibiting ability.

On repeat trials, the length of the phrase was the same as that on the previous trial, whereas the length of the phrases differed on switch trials. Thus, speakers could keep the previous task demands in working memory (e.g., to maintain and process information for two or three phrase components) on repeat trials, whereas they had to change the task demands on switch trials. This difference in processing demand was reflected in the RTs as the switch effect. In addition, there was an interaction with phrase length: The switch effect in RTs was present for the short phrases but not for the long ones. Individual differences in shifting scores, as obtained by the shape–colour switching task, did not correlate with the magnitude of the switch effect in the mean RTs. However, analyses of the shape of the RT distributions revealed that individual differences in shifting scores correlated positively with the switch effect in the normal part of the distribution and negatively with the switch effect in the tail. Thus, better shape–colour shifting scores were associated with a smaller switch effect on most of the picture-description trials but with a larger switch effect on the abnormally slow trials. These results suggest that phrase production is influenced by the shifting ability.

To summarize, we observed length, distractor, and switch effects in mean RTs and/or RT distributional components in noun-phrase production. The magnitude of these length, distractor, and switch effects correlated with individual differences in updating, inhibiting, and shifting scores, respectively. Thus, all three components of executive control distinguished by Miyake et al. (2000) appear to influence language production. Our findings are in agreement with previous studies that observed an influence of updating and inhibiting ability on picture naming (Piai & Roelofs, 2013; Shao et al., 2012). The present study not only shows that updating and inhibiting ability may have an influence on more complex forms of language production than picture naming—namely, phrase production—but also shows a novel influence of shifting ability. In the remainder of the manuscript, we discuss how the present findings relate to previous results in the literature and

what they imply for the role of executive control in language production.

Psycholinguistic findings

The aim of the present study was to examine the influence of executive control on noun-phrase production. To reduce error variance due to differences in processing speed among pictures and words, only a limited number of objects (four) and colours (two) were used in the present study. In contrast, previous investigations of phrase production in the literature used larger numbers of items. For example, in a seminal study of noun-phrase production, Schriefers (1992) used 32 pictures and four colours. Nevertheless, the present study replicated a number of basic psycholinguistic findings in the literature, including the effect of congruency of grammatical gender between picture name and distractor word, and the interaction between definiteness and gender congruency. We briefly discuss these findings in turn, starting with the basic effect of definiteness.

We observed that the picture description RTs were shorter for indefinite phrases than for definite phrases. For both phrase types, the grammatical gender of the noun needs to be processed. However, for definite phrases, the grammatical gender determines the article (*de* vs. *het*), whereas for indefinite phrases the gender determines the inflection of the adjective (e.g., *groene* vs. *groen*). In the production of indefinite phrases, inflectional encoding of the adjective may occur while the speaker is articulating the indefinite article (*een*, which is the same on all indefinite trials), whereas the definite article (*de* or *het*) needs to be encoded before articulation onset for the definite phrases. This factor may explain why articulation onset is somewhat later for definite phrases. However, although the definite and indefinite blocks of trials were alternated in the experiment (i.e., definite, indefinite, definite, indefinite, etc.), all participants received the blocks in the same order (cf. Friedman et al., 2008; Miyake et al., 2000), always starting with a definite block of trials. Thus, in performing the first indefinite block, participants had already responded to the pictures by producing definite phrases. Consequently, indefinite blocks may have been associated with a greater amount of practice with the pictures than definite blocks, yielding shorter RTs for the indefinite than for the definite condition. Thus, it remains possible that the finding of a shorter mean RT for the indefinite than for the definite condition is fully or partly due to a differential amount of practice.

We also observed a gender congruency effect, replicating Schriefers (1993), Schriefers and Teruel (2000), Schiller and Caramazza (2003), La Heij, Mak, Sander, and Willeboordse (1998), and Van Berkum (1997), among others. Picture description RTs were shorter when the incongruent distractor word had the same grammatical gender as the picture name than when it had a different gender. This gender congruency effect was obtained for the definite phrases but not for the indefinite phrases. This suggests that the gender congruency effect arose in selecting the appropriate article, which involved a choice for the definite phrases (between *de* or *het*) but not for the indefinite phrases (always *een*). Schiller and Caramazza (2003) obtained a similar finding for the production of noun phrases in Dutch and German. Participants had to refer to one or two pictured objects by producing singular or plural noun phrases. For the singular phrases, the article is gender-marked (*de* or *het* in Dutch), but for plural phrases it is not (i.e., *de* in Dutch regardless of grammatical gender). A gender congruency effect was obtained for the singular but not for the plural phrases, suggesting that the congruency effect arises in the selection of the article.

To conclude, although fewer pictures and colours were used in the present study than in previous studies in the literature, we replicated basic psycholinguistic findings. Thus, the influences of executive control on noun-phrase production in the present study were obtained in the context of a replication of basic psycholinguistic findings.

Updating ability and language production

We found that the length effect in phrase production RTs correlated with scores on the operation-span task but not with those on the odd-one-out task. Several studies in the literature suggest

that the operation-span and odd-one-out tasks both measure updating ability (Conway et al., 2005), whereas other studies make a distinction between verbal and nonverbal updating (e.g., Montgomery et al., 2010). We found that speakers with better verbal updating scores had a smaller length effect in the picture description RTs, whereas no such correlation was found with nonverbal updating scores. Thus, better verbal updating ability in particular may help to process and maintain a larger amount of information during language production. Moreover, both verbal and nonverbal updating scores correlated with the mean RTs across all conditions. This finding corresponds to the observations of Shao et al. (2012) and Piai and Roelofs (2013), showing that verbal updating scores correlate with picture-naming RT. In line with the present study, speakers with better verbal updating scores were faster to name pictures in those previous studies. We not only replicated this correlation between verbal updating scores and RT for the production of noun phrases, but also observed that this correlation is obtained for both verbal and nonverbal updating scores.

It is not clear why verbal but not nonverbal updating scores correlated with the length effect in the RTs, whereas both scores correlated with overall mean RT. This difference in correlation pattern may be further investigated in future research. Regardless, the key point here is that our results indicate that updating ability influences phrase production.

We also found correlations between the scores of both the odd-one-out and operation-span tasks and the mean of the normal part of the RT distribution (i.e., the μ parameter of the ex-Gaussian distribution) across all conditions. As discussed before, this μ parameter represents the majority of the trials making up the distribution. These findings suggest that verbal and nonverbal updating abilities influence the speed of phrase production on a regular basis. Speakers with better updating scores were able to describe the pictures faster on most of the trials. In contrast, Shao et al. (2012) found that the (verbal) updating scores correlated with the τ parameter of the RT distribution of picture naming. The τ parameter represents the right tail of the RT distribution, made up by the abnormally slow responses, suggesting that, in picture naming, updating is reflected only in the slowest responses. In contrast, in the present study we found that updating was related to the normal part rather than the tail of the RT distribution. The phrase production task used in our study is more complex than picture naming. In more demanding situations like phrase production, speakers need to plan and process more information. Therefore, updating ability may be more regularly engaged in picture description requiring phrase production than it is in picture naming involving only single word production. Consequently, individual differences in updating ability are reflected in the normal part of the RT distribution in phrase production but only in the distribution tail in picture naming, as empirically observed.

To summarize, we found that both verbal and nonverbal updating abilities influenced the overall speed of phrase production. Individuals with better updating ability produced the phrases faster on most of the trials. Moreover, we observed that better verbal updating ability was associated with a smaller magnitude of the length effect. Speakers with better updating ability had a smaller difference in RT between the long and small phrases. These results provide evidence that updating ability is important for phrase production.

Inhibiting ability and language production

We found a correlation between the scores for the stop-signal task and the distractor effect in the picture description RTs. Individuals with better inhibiting scores had smaller distractor effects in the RTs. Thus, inhibiting ability may influence how well speakers can overcome interference from irrelevant spoken information during phrase production. It is generally assumed that a picture of an object activates not only its name in lexical memory but also, to a lesser extent, the names of semantically related objects (e.g., Levelt et al., 1999; Piai et al., 2014; Roelofs, 1992, 2003). For example, a picture of a fork activates not only the word *fork* but also competing words such as *plate*, *bottle*, and *glass*. An incongruent spoken distractor

(e.g., *plate*) will boost the activation of one of those competitors, whereas a congruent distractor will activate the picture name. The present results suggest that speakers with better inhibiting abilities are able to suppress the competitors better than speakers with poorer inhibiting abilities.

Similar to Shao et al. (2012), we did not find a correlation between the updating and inhibiting scores. Moreover, updating scores were correlated with the length effect in RTs, whereas the inhibiting scores were correlated with the distractor effect. As suggested by Miyake et al. (2000), inhibiting and updating abilities are separable components of executive control. Our data suggest that, at least to some extent, the influences of updating and inhibiting abilities on language production are also separable.

In addition, we investigated what kind of inhibition influences language production. We obtained evidence suggesting that both selective and nonselective inhibition are involved. First, we found a correlation between the magnitude of the distractor effect in the mean picture description RTs and the scores on the stop-signal task, which are taken to index nonselective inhibiting ability (Forstmann et al., 2008). Second, we found a correlation between the magnitude of the distractor effect in the mean picture description RTs and the slopes of the slowest delta segment, which are taken to index selective inhibiting ability (Forstmann et al., 2008). We did not find a correlation between the scores on the stop-signal task and the slopes of the slowest delta segment, which suggests that these measures index separable kinds of inhibiting abilities (i.e., nonselective and selective).

Whereas Shao et al. (2012) found a correlation between the scores on the stop-signal task and picture-naming RT, we did not observe such a correlation. In the study of Shao et al., no distractors were presented. Thus, the inhibiting ability influenced the degree to which competitors activated by the picture affected the naming RT. In the present study, a congruent or incongruent spoken distractor was presented on all trials, which dominated the competition on each trial. Consequently, the influence of inhibiting ability was reflected in the correlation with the magnitude of the distractor effect rather than the picture description RT per se.

To summarize, we obtained evidence for the involvement of both selective and nonselective inhibiting in phrase production. Our findings suggest that speakers with better selective as well as nonselective inhibiting abilities had smaller distractor effects in the picture description RTs. These results provide evidence that inhibiting ability is important for phrase production.

Shifting ability and language production

We did not obtain a correlation between shifting scores and the magnitude of the switch effect in the picture description RTs. However, RT distribution analyses revealed that shifting scores correlated positively with the switch effect in the normal part of the distribution and negatively with the switch effect in the distribution tail. Thus, better shifting scores were associated with a smaller switch effect on most of the trials but with a larger switch effect on the abnormally slow trials. Similar to the findings of Shao et al. (2012) for picture naming, we did not find a correlation between shifting scores and the overall mean picture description RTs. However, the role of shifting ability in simple picture naming is unclear, whereas it is much clearer in a situation involving actual switching, such as in the present study. This explains why we did obtain an influence of shifting ability, whereas Shao et al. did not.

We found a positive correlation between the shifting scores and the magnitude of the switch effect in the normal part of the RT distribution and a negative correlation with the switch effect in the tail. These opposite correlations explain why no overall correlation was found for the switch effect in the mean RT. The positive correlation between the shifting scores and the switch effect in the normal part of the distribution is as expected. Speakers with better shifting ability tend to have smaller switch effects on most of the trials. However, the negative correlation between the shifting scores and the switch effect in the tail of the distribution is unexpected. It implies that speakers with better shifting ability tend to have

larger rather than smaller switch effects on the trials associated with long responses. This surprising finding may be due to the following. Participants had a limited amount of time to complete each trial. On trials that were for some reason difficult (for instance, because of a lapse of attention), speakers with poorer shifting ability might not have been able to complete the noun phrase in time, resulting in an error rather than a very slow response. In contrast, speakers with better shifting ability might have been able to complete the noun phrase in time, resulting in a very slow response rather than an error. As a consequence, the magnitude of the switch effect in the slowest responses may be larger for speakers with better than for those with poorer shifting ability, as we observed.

We suggest that speakers with better shifting ability would tend to be able to complete the phrase in time on trials that were for some reason difficult (the slowest trials making up the tail of the RT distribution), whereas speakers with poorer shifting ability would not tend to complete a correct phrase in time on those difficult trials. This would result in a larger magnitude of the switch effect in the slowest responses (i.e., the τ parameter) for speakers with better shifting ability. Thus, speakers with poorer shifting ability are expected to have more time-out errors on switch trials, predicting a positive correlation between the shifting ability as measured by shape–colour task and the switch effect on time-out errors. We calculated the difference in the number of time-out errors between switch and repeat trials, and we found that it correlated positively with the scores on the shape–colour task, $r = .299$, $p < .05$. This suggests that subjects with poorer updating ability had indeed more time-out errors on switch than on repeat trials when compared with subjects with better updating ability. For subjects making time-out errors, we assessed the correlation between the shape–colour task scores and the switch effect in τ and found a significant negative correlation, $r = -.612$, $p < .05$. Thus, poor shifting ability leads to a larger switch effect on time-out errors and a smaller switch effect in τ. This supports our conjecture that on difficult trials, speakers with better shifting ability are able to complete phrases in time, whereas speakers with poorer shifting ability are not.

We obtained an interaction between switching and length: The switch effect in mean RTs was obtained for short phrases but not for long phrases. This observation corresponds to the asymmetry in switch costs that is often obtained in task switching. For example, in switching between Stroop colour naming and word reading, a switch effect is obtained for the (faster) word reading task but not for the (slower) colour naming task (e.g., Allport & Wylie, 1999, 2000; Gilbert & Shallice, 2002). Similarly, in the present study, a switch effect was obtained for the (faster) short phrase production but not for the (slower) long phrase production. According to Allport and Wylie (1999, 2000), the switch effect is caused by task-set inertia: The irrelevant task set of the previous trial is still active on the current trial and needs to be actively disengaged. Asymmetrical switch effects may occur when the disengagement takes longer for one task than the another because of differential positive or negative priming of task set. For example, word reading is faster than colour naming. Therefore, to prevent inadvertent word reading responses on Stroop colour naming trials, the task set of the word reading task may be inhibited, and the task set of colour naming may be enhanced. In contrast, on word reading trials, inhibition of colour naming and enhancement of word reading are not needed. As a consequence, disengagement from the previous task set will take much longer in switching to word reading than to colour naming. Switching to word reading will be delayed because of positive priming of task set (i.e., previous enhancement of the colour naming task), negative priming of task set (i.e., previous inhibition of the word reading task), or both.

In the present study, picture description RTs were shorter for the short than for the long phrases. Coloured pictures (e.g., a green fork) allowed both long-phrase responses ("de groene vork") as well as short-phrase responses ("de vork"), whereas the noncoloured pictures allowed for short-phrase responses only (i.e., "de groene vork" cannot be produced in response to a fork in

white colour). Therefore, to prevent the inadvertent production of short phrases in response to coloured pictures, the task set for producing short phrases may be inhibited, and the task set of producing long phrases may be enhanced. In contrast, on trials with noncoloured pictures requiring a short phrase, inhibition of long phrases or enhancement of short phrases is not needed (because the pictures only allow a short-phrase response). As a consequence, disengagement from the previous task set will take much longer in switching to a short phrase than to a long phrase, as we observed. Switching to the production of a short phrase will be delayed because of positive priming of task set (i.e., previous enhancement of the task set for producing a long phrase), negative priming of task set (i.e., previous inhibition of the task set for producing a short phrase), or both.

The switching between phrase types is assumed to engage a shifting ability that is domain general, as is generally assumed in the executive control literature (e.g., Miyake et al., 2000). It now seems generally accepted in the task switching literature that asymmetrical switch costs are due to differential task-set inertia (Koch, Gade, Schuch, & Philipp, 2010), as originally proposed by Allport and colleagues (e.g., Allport & Wylie, 1999, 2000). This idea is also accepted by Monsell and colleagues (Yeung & Monsell, 2003), who originally proposed an alternative account. Moreover, also in the bilingual literature, language switching is assumed to engage a domain-general shifting ability (e.g., Abutalebi & Green, 2007). The task-set inertia idea seems to be generally accepted as an account of asymmetrical language switch costs (e.g., Koch et al., 2010, for a review), following the seminal proposal by Meuter and Allport (1999).

To conclude, the interaction between switch and length effects in the RTs in the present study is in line with previous findings obtained for switching between tasks of different strengths. The present interaction may receive an analogous explanation in terms of task set inertia (positive or negative priming of the previous task set), which may be different for the production of short and long phrases.

Conclusion

The goal of the present study was to examine whether the updating, inhibiting, and shifting abilities underlying executive control influence spoken noun-phrase production. We obtained evidence that all three executive abilities influence noun-phrase production and that they contribute to varying extents and in different ways.

REFERENCES

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*, 242–275.

Allport, A., & Wylie, G. (1999). Task-switching: Positive and negative priming of task-set. In G. W. Humphreys, J. Duncan, & A. Treisman (Eds.), *Attention, space, and action: Studies in cognitive neuroscience* (pp. 273–296). Oxford: Oxford University Press.

Allport, A., & Wylie, G. (2000). 'Task-switching', stimulus-response bindings, and negative priming. In S. Monsell & J. S. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 35–70). Cambridge, MA: MIT Press.

Altmann, L. J. P., & Troche, M. S. (2011). High-level language production in Parkinson's disease: A review. *Parkinson's Disease*, Article ID 238956.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*(4), 170–177.

Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology Section A*, *49A*, 5–28.

Barkley, R. A. (2012). *Executive functions: What they are, how they work, and why they evolved*. New York, NY: Guilford Press.

Bastiaanse, R., & Leenders, K. L. (2009). Language and Parkinson's disease. *Cortex*, *45*, 912–914.

Brown, S., & Heathcote, A. (2003). QMLE: Fast, robust, and efficient estimation of distribution functions based on quantiles. *Behavior Research Methods, Instruments, & Computers*, *35*, 485–492.

Brownsett, S. L. E., Warren, J. E., Geranmayeh, F., Woodhead, Z., Leech, R., & Wise, R. J. S. (2014). Cognitive control and its impact on recovery from aphasic stroke. *Brain*, *137*, 242–254.

Bruce, B., Thernlund, G., & Nettelbladt, U. (2006). ADHD and language impairment: A study of the parent questionnaire FTF (Five to Fifteen). *European Child & Adolescent Psychiatry*, *15*(1), 52–60.

Collette, F., Van der Linden, M., Laureys, S., Delfiore, G., Degueldre, C., Luxen, A., & Salmon, E. (2005). Exploring the unity and diversity of the neural substrates of executive functioning. *Human Brain Mapping*, *25*, 409–423.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769–786.

Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 170–189.

Forstmann, B. U., Jahfari, S., Scholte, H. S., Wolfensteller, U., van den Wildenberg, W. P. M., & Ridderinkhof, K. R. (2008). Function and structure of the right inferior frontal cortex predict individual differences in response inhibition: A model-based approach. *The Journal of Neuroscience*, *28*, 9790–9796.

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*, 201–225.

Gilbert, S. J., & Burgess, P. W. (2008). Executive function. *Current Biology*, *18*, R110–R114.

Gilbert, S. J., & Shallice, T. (2002). Task switching: A PDP model. *Cognitive Psychology*, *44*, 297–337.

Goldstein, S., & Naglieri, J. A. (Eds.). (2014). *Handbook of executive functioning*. New York, NY: Springer.

Koch, I., Gade, M., Schuch, S., & Philipp, A. M. (2010). The role of task inhibition in task switching: A review. *Psychonomic Bulletin and Review*, *17*, 1–14.

Korvorst, M., Roelofs, A., & Levelt, W. J. M. (2006). Incrementality in naming and reading complex numerals: Evidence from eyetracking. *The Quarterly Journal of Experimental Psychology*, *59*, 296–311.

La Heij, W., Mak, P., Sander, J., & Willeboordse, E. (1998). The gender-congruency effect in picture-word tasks. *Psychological Research*, *61*(3), 209–219.

Leonard, L., Ellis Weismer, S., Miller, C., Francis, D., Tomblin, J. B., & Kail, R. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research*, *50*, 408–428.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain sciences*, *22*(1), 1–38.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, *98*, 13464–13471.

Logan, G. D. (1985). Executive control of thought and action. *Acta Psychologica*, *60*, 193–210.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York, NY: Oxford University Press.

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*, 25–40.

Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, *48*, 131–147.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review*, *104*, 3–65.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140.

Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working memory and specific language impairment: An update on the relation and perspectives on assessment and treatment. *American Journal of Speech-Language Pathology*, *19*(1), 78–94.

Piai, V., & Roelofs, A. (2013). Working memory capacity and dual-task interference in picture naming. *Acta Psychologica*, *142*, 332–342.

Piai, V., Roelofs, A., Acheson, D. J., & Takashima, A. (2013). Attention for speaking: Domain-general control from the anterior cingulate cortex in spoken word production. *Frontiers in Human Neuroscience*, *7*, 1–14.

Piai, V., Roelofs, A., Jensen, O., Schoffelen, J.-M., & Bonnefond, M. (2014). Distinct patterns of brain activity characterise lexical activation and competition in spoken word production. *PLOS ONE*, *9*, e88674, 1–11.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, *86*, 446–461.

Ridderinkhof, K. R. (2002). Activation and suppression in conflict tasks: Empirical clarification through distributional analyses. In W. Prinz & B. Hommel (Eds.), *Attention and performance XIX: Common mechanisms in perception and action* (pp. 494–519). Oxford: Oxford University Press.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*, 107–142.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, *110*, 88–125.

Roelofs, A. (2004). Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: Comment on Rapp and Goldrick (2000). *Psychological Review*, *111*, 561–572.

Roelofs, A. (2006). Context effects of pictures and words in naming objects, reading words, and generating simple phrases. *The Quarterly Journal of Experimental Psychology*, *59*, 1764–1784.

Roelofs, A., & Piai, V. (2011). Attention demands of spoken word planning: A review. *Frontiers in Psychology*, *2*, 1–14.

Schiller, N. O., & Caramazza, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language*, *48*, 169–194.

Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., & Thompson-Schill, S. L. (2009). Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proceedings of the National Academy of Sciences*, *106*(1), 322–327.

Schriefers, H. (1992). Lexical access in the production of noun phrases. *Cognition*, *45*, 33–54.

Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 841–850.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, *29*(1), 86–102.

Schriefers, H., & Teruel, E. (2000). Grammatical gender in noun phrase production: The gender interference effect in German. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1368–1377.

Shao, Z., Meyer, A. S., & Roelofs, A. (2013). Selective and nonselective inhibition of competitors in picture naming. *Memory & Cognition*, *41*, 1200–1211.

Shao, Z., Roelofs, A., & Meyer, A. S. (2012). Sources of individual differences in the speed of naming objects and actions: The contribution of executive control. *The Quarterly Journal of Experimental Psychology*, *65*, 1927–1944.

Van Berkum, J. J. A. (1997). Syntactic processes in speech production: The retrieval of grammatical gender. *Cognition*, *64*(2), 115–152.

Van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavioural Research Methods*, *38*(4), 584–589.

Van den Wildenberg, W. P. M., Wylie, S. A., Forstmann, B. U., Burle, B., Hasbroucq, T., & Ridderinkhof, R. K. (2010). To head or to heed? Beyond the surface of selective action inhibition: A review. *Frontiers in Human Neuroscience*, *4*, 1–13.

Verbruggen, F., Logan, G. D., & Stevens, M. A. (2008). STOP-IT: Windows executable software for the stop signal paradigm. *Behavior Research Methods*, *40*, 479–483.

Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition*, *75*, 105–143.

Yeung, N., & Monsell, S. (2003). Switching between tasks of unequal familiarity: The role of stimulus-attribute and response-set selection. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 455–469.