THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2006, 59 (2), 296–311

Psychology Press
Taylor & Francis Group

Incrementality in naming and reading complex numerals: Evidence from eyetracking

Marjolein Korvorst
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands and University Hospital RWTH Aachen, Aachen, Germany

Ardi Roelofs
Max Planck Institute for Psycholinguistics, Nijmegen, F. C. Donders Centre for Cognitive Neuroimaging, Nijmegen, and Nijmegen Institute for Cognition and Information, Nijmegen, The Netherlands

Willem J. M. Levelt
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Individuals speak incrementally when they interleave planning and articulation. Eyetracking, along with the measurement of speech onset latencies, can be used to gain more insight into the degree of incrementality adopted by speakers. In the current article, two eyetracking experiments are reported in which pairs of complex numerals were named (arabic format, Experiment 1) or read aloud (alphabetic format, Experiment 2) as house numbers and as clock times. We examined whether the degree of incrementality is differentially influenced by the production task (naming vs. reading) and mode (house numbers vs. clock time expressions), by comparing gaze durations and speech onset latencies. In both tasks and modes, dissociations were obtained between speech onset latencies (reflecting articulation) and gaze durations (reflecting planning), indicating incrementality. Furthermore, whereas none of the factors that determined gaze durations were reflected in the reading and naming latencies for the house numbers, the dissociation between gaze durations and response latencies for the clock times concerned mainly numeral length in both tasks. These results suggest that the degree of incrementality is influenced by the type of utterance (house number vs. clock time) rather than by task (reading vs. naming). The results highlight the importance of the utterance structure in determining the degree of incrementality.

Individuals speak incrementally when they interleave planning and articulation, while describing scenes, naming objects, reading aloud, and so on. Eyetracking, along with the measurement of speech onset latencies, can be used to gain more insight into the degree of incrementality adopted by speakers. Previous studies of object naming have suggested that there is a tight link between eye movements and speech-planning processes (e.g., Griffin & Bock, 2000; Meyer & Van der Meulen, 2000). In particular, gaze durations in object naming have been shown to depend on

---

Correspondence should be addressed to Marjolein Korvorst (née Meeuwissen), University Hospital RWTH Aachen, Department of Neurology, Section Neuropsychology, Pauwelsstrasse 30, D-52074 Aachen, Germany. Email: korvorst@neuropsych.rwth-aachen.de

This research was made possible by a PhD stipend from the Max Planck Society to the first author. We wish to thank Kay Bock, Marc Brysbaert, Marco Zorzi, the members of the Utterance Encoding Group at the Max Planck Institute for Psycholinguistics in Nijmegen, and one anonymous reviewer for helpful comments and discussions.

DOI:10.1080/17470210500151691

the time to plan the sound form of the corresponding name (Griffin, 2001; Meyer, Sleiderink, & Levelt, 1998). For example, when speakers have to first name a left object and then a right object, they look longer at left objects with two- than with one-syllable names even when the object recognition times are the same (Meyer, Roelofs, & Levelt, 2003). Furthermore, speakers look longer at objects with low-frequency than with high-frequency names (Griffin, 2001; Meyer et al., 1998). Moreover, a study by Meyer and Van der Meulen (2000) showed that response latencies and gaze durations in object naming are affected in a similar way by auditory primes. Speakers were presented with object pairs together with auditory prime words, which could be either phonologically related or unrelated to the name of the left object. Both response latencies and gaze durations for the left object were shorter with related than with unrelated auditory primes. This held regardless of whether the auditory words primed the first or the second syllable of the object name. These effects of word length, frequency, and phonological priming indicate that the shift of gaze from the left to the right object is initiated only after the sound form of the name for the left object has been planned. Thus, gaze durations can provide a good measure for estimating the total amount of speech planning that is required in order to produce an utterance when describing an object (cf. Meyer & Lethaus, 2004, for a review of the literature).

Gaze durations and naming latencies do not always point in the same direction, however, as shown by Levelt and Meyer (2000). They reported that response latencies and gaze durations can dissociate in that gaze durations may reflect the phonological length (e.g., number of syllables) of the utterance even when response latencies do not. In their study, speakers were instructed to describe coloured left and right objects (e.g., a big red scooter and a ball) in a simple or a complex way. Participants had to respond with either "the scooter and the ball" or "the *big red* scooter and the ball". Gaze durations for the left object (the scooter) were much shorter for the simple utterances (559 ms) than for the complex utterances (1,229 ms). However, speech onset latencies did not differ between the two utterance types. Furthermore, the shift of gaze to the right object was initiated before speech onset for the simple utterances, but after speech onset for the complex utterances. This suggests that the shift of gaze, but not the onset of articulation, is triggered by the completion of phonological encoding of the utterance referring to the first object. It seems that in producing the long utterances (e.g., "the *big red* scooter and the ball"), the participants adopted an incremental approach to speech production—that is, they simultaneously articulated (as revealed by speech onset latencies) and planned (as revealed by gaze durations) the utterances.

The issue of incrementality in spoken-language production was recently also examined by Ferreira and Swets (2002). In their experiments, participants stated the (two-digit) sums of addition problems, using three different types of utterance forms: (a) the answer itself; (b) the answer at the beginning of a sentence ("... is the answer"); and (c) the answer at the end of a sentence ("the answer is ..."). Note that when the answer was at the end, speakers would be able to carry out some of the planning of the sum while already articulating the beginning of the sentence. The addition problems differed in the difficulty with which the sum could be calculated. Moreover, participants were given ample time for responding (Experiment 1) or they were put under time pressure by giving them a response deadline (Experiment 2). In Experiment 1, speech onset latencies reflected the problem difficulty but not the utterance type. This suggests that the participants initiated articulation only after they had calculated the sum, regardless of utterance type. Surprisingly, the same results were obtained in Experiment 2, when participants were put under time pressure. Some evidence for incrementality was obtained, though, in articulation durations. The articulation durations varied with problem difficulty, which suggests that some planning occurred while articulating. According to Ferreira and Swets, these results suggest that the extent to which people speak incrementally is under strategic control—for example, an incremental approach is more likely to be adopted by speakers when put under time pressure. Similarly, Levelt and Meyer (2000) suggested that incrementality is more likely to occur in speaking when long utterances have to be produced.

In the current article, we report two eyetracking experiments that further explored the issue of incrementality in speech production. The results of Levelt and Meyer (2000) suggest that incrementality is likely to occur when long utterances have to be produced in a speeded task. We tried to optimize the chance that individuals adopted an incremental approach by using complex numerals consisting of three digits, as previously done by Meeuwissen, Roelofs, and Levelt (2003), rather than the two-digit numerals used in the study by Ferreira and Swets (2002). Moreover, participants had to produce utterances that included two such complex numerals rather than one. As a result, quite long utterances had to be produced: for instance, "two hundred forty-five and three hundred fifty-five". In order to remain fluent, speakers might choose to adopt an incremental approach by interleaving their planning and speaking. In turn, this might result in speech being initiated relatively early in time by speakers. If that is the case, speech onset latencies might not be the best dependent variable for measuring the amount of speech planning. Therefore, we decided to monitor both gaze durations and speech onset latencies in the current experiments.

Earlier research on spoken complex numeral production in Dutch has provided evidence about the speech-planning variables that are involved in the naming and reading of complex numeral expressions. Two types of responding to an identical complex numeral can be compared by contrasting house numbers and clock times (e.g., 235 vs. 2:35). House numbers and clock times differ in the utterance structure and the planning levels engaged. Meeuwissen et al. (2003) obtained evidence for equivalent speech-planning processes being involved in naming house numbers from arabic digit format (e.g., 235) and reading them aloud from alphabetic format (e.g., Dutch TWEEHONDERDVIJFENDERTIG). For both naming and reading alike, response latencies were determined by morphophonological factors, such as the number of phonemes and morphemes, and the log morpheme and whole-form frequencies. For clock time naming (i.e., arabic format), however, response latencies not only demonstrated the influence of morphophonological factors, but also of conceptual factors, reflecting the conceptual operations required in telling time in the Dutch language (see also Meeuwissen, Roelofs, & Levelt, 2004, 2005). In Dutch, time is told in a relative way, with time expressions not only making reference to the full hour, like in English, but also to the half hour (Bock, Irwin, Davidson, & Levelt, 2003). This secondary referent in Dutch applies between 10 minutes before and 10 minutes after the half hour, yielding, for example, "tien voor half drie" (literal translation: "ten before half three") for 2:20. To tell time from a digital clock in Dutch, the digits in the input have to be used to determine the utterance referent and the distance in minutes from that referent. These conceptual operations were reflected in the naming latencies of the clock times. When the conceptual operations were made unnecessary by presenting the clock times in alphabetic format, the response latencies were only determined by morphophonological factors, just as they were in reading the house numbers.

These previous experiments measured naming and reading latencies but not gaze durations. Other studies that investigated numeral naming recorded gaze durations but not naming latencies. For instance, Pynte (1974) measured eye movements during the silent naming of two-digit arabic numerals and observed that gaze durations increased with the number of syllables in the numeral names. Pynte's stimuli differed not only in number of syllables but also in number magnitude and numeral frequency. In order to determine the contribution of all three factors, Gielen, Brysbaert, and Dhondt (1991) conducted multiple regression analyses on the gaze durations in the silent naming of all one- and two-digit arabic numerals ranging from 0 to 99. The regressions indicated that the number of syllables and the numeral frequency, but not the number magnitude, determined the gaze durations, even when the contribution of the other factors was partialled out. However, the results by Gielen et al. (1991) were contested in a follow-up study by Brysbaert (1995). When controlling for several methodological factors, such as lack of power and inclusion of the numeral 0, Brysbaert (1995) showed that gaze durations for numerals ranging from 1 to 99 were a function of the logarithm of magnitude.

Below, we report two experiments that examined the naming and oral reading of complex numeral pairs while measuring both response latencies and gaze durations. The aim was to examine whether the extent to which individuals articulate and plan simultaneously is influenced by the production task (naming vs. reading) and the utterance mode (house number vs. clock time expression). The pairs of complex numerals were named (arabic format, Experiment 1) or read aloud (alphabetic format, Experiment 2) as house numbers and as clock times. Comparing reading with naming should reveal whether participants are more likely to articulate while planning, as revealed by dissociations between speech onset latencies (reflecting articulation) and gaze durations (reflecting planning), in reading than in naming the complex numeral pairs. Throughout the article, the terms "reading" and "naming" refer to the difference in input, alphabetic versus arabic digit, respectively. Moreover, comparing house numbers with clock times should reveal whether participants are more likely to articulate while planning when producing utterances that require the involvement of fewer speech-planning stages for successful production, which is the case in producing house number pairs (which only require morphophonological speech planning) as compared with producing clock time pairs (which require both conceptual and morphophonological speech planning), see Meeuwissen et al. (2003, 2005).

EXPERIMENT 1

In the first experiment, speakers named complex numeral pairs presented in arabic digit format as house numbers or as clock times. We measured gaze durations and naming latencies for the first numeral. We expected that the gaze durations would mirror the naming latency results found earlier, with morphophonological factors determining the naming of house numbers, and both morphophonological and conceptual factors determining the naming of clock times. The first question was whether dissociations would be obtained between naming latencies and gaze durations in that gaze durations reflect the morphophonological and conceptual factors even when naming latencies do not. Such dissociations would provide evidence for incrementality in naming complex numerals. The second question was whether the dissociations would depend on the utterance mode—that is, whether the dissociations would occur for both the house numbers and the clock times.

Method

Participants

A total of 20 native speakers of Dutch participated in the experiment. They were undergraduate students of Nijmegen University. They had normal or corrected-to-normal vision. They were paid for their participation.

Materials and design

Stimuli for both the house number and the clock time mode (96 in total) consisted of complex arabic numerals ranging from 200 to 955 and ending on either 0 or 5. This yielded 12 different number types collapsed across the hundreds and the hours: (00) 200, 300, ... ; (05) 205, 305, ... ; (10) 210, 310, ... ; (15) 215, 315, ... ; (20) 220, 320, ... ; (25) 225, 325, ... ; (30) 230, 330, ... ; (35) 235, 335, ... ; (40) 240, 340, ... ; (45) 245, 345, ... ; (50) 250, 350, ... ; (55) 255, 355, ... For each type, there were eight instances—for example, for type (00), the instances were 200, 300, 400, 500, 600, 700, 800, and 900. Note that complex numerals ranging from 100 to 155 were excluded from the stimulus set. For numbers in the interval of 100 to 199, in Dutch no explicit information is encoded about the particular hundred involved (deletion of "one" for the first hundred is also common in French, cf. Barrouillet, Camos, Perruchet, & Seron, 2004). For instance, in Dutch, the number 105 is pronounced as "honderdvijf" (literal translation: "hundred five"), whereas the number 205 is pronounced as "*twee*honderdvijf" (literal translation: "*two* hundred five").

Each numeral that was presented on the left half of the screen occurred in three different pairs. A particular pair of numerals occurred only once in the whole experiment (i.e., irrespective of the left- or right-hand position of the stimuli), and each left-hand numeral was paired with numerals that belonged to a different group of hundreds (e.g., 200 and 305 together, but not 200 and 245). Numerals were scaled to fit into a virtual frame of 8.18 by 7.45 cm, corresponding to visual angles of 7.1° horizontally and 6.5° vertically when viewed from the participant's position, approximately 65 cm away from the screen. The distance between the midpoints of these virtual frames was 14.97 cm (13°).

Each mode of response (house number, clock time) was tested in a separate block of 288 (12 types × 8 instances × 3 pairings) trials. The order of presenting the numeral pairs was random. The order of testing each response mode was counterbalanced across participants.

Apparatus

A Compaq 486 computer controlled the experiment. Materials were presented on a ViewSonic 17PS screen. Eye movements were measured using an SMI EyeLink-Hispeed 2D head-mounted eyetracking system (SensoMotoric Instruments GmbH, Teltow, Germany). The position of the right eye was determined every 4 ms. Participants' speech was recorded using a Sennheiser ME400 microphone and a SONY DTC55 DAT recorder. Response latencies were measured using a voice key.

Procedure

The participants were tested individually. They were told that they would see pairs of complex numerals, which they should name in noun phrases either as house numbers or as clock times, with the conjunction *en* ("and") between the numerals. Each response had to start with the same word: *op* (literal translation: "at") for the house number mode and *om* (literal translation: "at") for the clock time mode. In this way, all responses started with the same phoneme. Importantly, participants were instructed to respond in a fluent manner, such that there were no pauses between the words *op*/*om* (at) and the complex numeral pair (house number or clock time). Trials on which this was not the case were excluded from analyses.

The headband of the eyetracking system was placed on the participant's head, and the system was calibrated. For calibration, a grid of three by three positions had been defined. During a calibration trial, a fixation target appeared once, in random order, in each of these positions for one second. Participants were asked to fixate upon each target until the next target appeared. After the calibration trial, the estimated positions of the participant's fixations and the distances from the fixation targets were displayed to the experimenter. Calibration was considered adequate if there was at least one fixation within 1.5° of each fixation target. When calibration was inadequate, the procedure was repeated, sometimes after adjusting the eye cameras. Successful calibration was followed by a validation trial. For the participants, this trial did not differ from the calibration trial, but the data collected during the validation trial were used to estimate the participants' gaze positions, and the error (i.e., the distance between the estimated gaze position and the target position) was measured. Validation was considered completed if the average error was below 1 degree and the worst error below 1.5 degrees. Depending on the result of the validation trial, the calibration and validation trials were repeated or testing began. Calibration was repeated before the experiment and during the breaks.

A fixation point appeared in the centre of the frame for the left complex numeral in the pair for 850 ms at the beginning of each test trial. Previous studies by Meyer et al. (1998, 2003) showed that participants have a strong tendency to inspect and name the left stimulus first, which was reinforced here by the presentation of the fixation point. Next, the numeral pair was presented and remained on the screen until the participant had fully completed the task and had pushed a button for the next trial. After every 48 trials there was a short break.

Analyses. A trial was considered invalid when it included a speech error, when a wrong oral response was produced, when the oral response was not fluent, or when the voice key was triggered incorrectly. Error trials were discarded from the analyses of gaze durations and response latencies. To analyse the speakers' gaze durations, we first classified their eye fixations as falling on the left or right numeral or elsewhere. A fixation was counted as on a numeral if its coordinates lay within or on the outer contours of the numeral. Furthermore, trials on which speakers did not fixate on the left numeral first before turning to the right numeral were excluded from analyses (less than 1% of all cases). The first-pass gaze duration was defined as the time interval between the beginning of the first fixation on the left-hand numeral and the end of the last fixation before the first shift of gaze was initiated to the right-hand numeral.

The naming latencies, gaze durations, and errors were submitted to analyses of variance (ANOVAs), with the crossed variables mode and type. For these analyses, we report *minF′* statistics (Clark, 1973; Raaijmakers, Schrijnemakers, & Gremmen, 1999). Interactions of mode and type were further explored through multilevel multiple regressions on the total data set. For all analyses (ANOVA and multilevel multiple regression) reported in this article, an alpha level of .05 was adopted. For all best fitting regression models reported in this paper adding another group of factors did not significantly increase the amount of variance accounted for.

Results

Table 1 gives the mean naming latencies, their standard deviations, the first-pass gaze durations,

Table 1. *Mean latencies*[a] *for the vocal responses, gaze shifts*[a], *and error percentages, per mode and type for Experiment 1*

	Mode									
	House number					*Clock time*				
	Vocal		*Gaze*			*Vocal*		*Gaze*		
Type	*M*	*SD*	*M*	*SD*	*% Error*	*M*	*SD*	*M*	*SD*	*% Error*
00	545	190	749	332	3.3	615	195	633	212	5.6
05	545	206	925	326	4.0	641	212	967	324	7.3
10	551	196	889	322	3.7	660	213	919	311	5.6
15	549	201	1,007	336	5.2	660	212	952	307	8.1
20	553	197	976	334	3.5	742	269	1,322	453	7.7
25	549	191	1,172	364	5.2	726	252	1,300	432	7.1
30	551	197	960	340	5.4	703	249	920	310	8.8
35	555	208	1,227	384	6.5	702	249	1,367	416	6.3
40	556	203	1,001	365	5.8	731	246	1,375	457	8.5
45	562	208	1,274	397	5.6	700	231	997	361	7.5
50	558	214	1,023	476	3.3	722	274	1,111	434	9.0
55	564	208	1,245	387	4.2	679	253	1,051	357	7.7
All types	553	202	1,037	397	4.7	689	241	1,076	430	7.4

[a] In ms.

their standard deviations, and the error percentages for the 12 numeral types in the house number and clock time modes. Figure 1 displays the mean gaze durations and naming latencies for the house numbers (upper panel) and clock times (lower panel).

Errors

Participants made fewer errors in the house number mode than in the clock time mode, *minF*′(1, 69) = 15.32, $p < .01$. As indicated by Table 1, most errors were made in the slowest condition, so there is no evidence for a speed–accuracy tradeoff. There were no other significant effects.

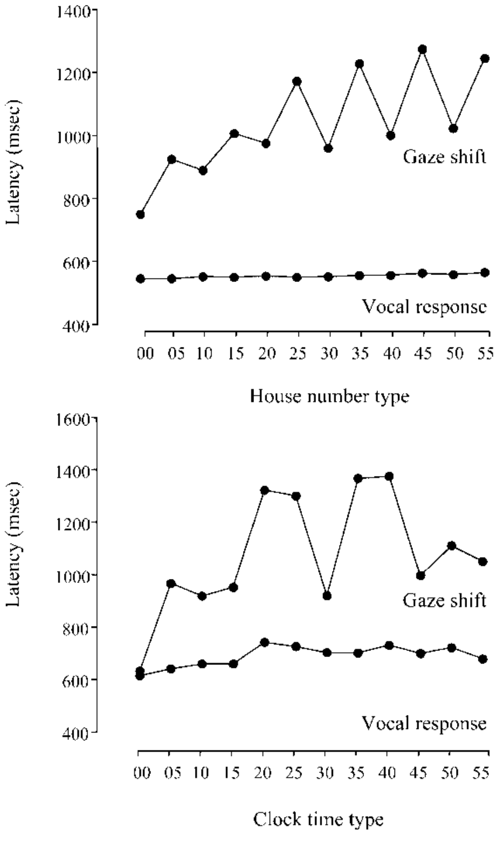

Figure 1. *Mean latencies of the vocal response and the shift of gaze in naming house numbers and clock times in Experiment 1.*

Naming latencies

Naming latencies were on average 136 ms shorter for the house numbers than for the clock times, *minF*′(1, 20) = 26.57, $p < .01$. Moreover, naming latencies depended on type, *minF*′(11, 270) = 4.68, $p < .01$. Furthermore, the effect of type varied with mode, *minF*′(11, 212) = 3.36, $p < .01$.

To gain further insight into the factors underlying the interaction between type and mode, we ran multiple regression analyses over the total data set. The predictor variables concerned either conceptual or morphophonological factors involved in speech planning. The following groups of predictor variables were entered into the analyses for conceptual level planning: (a) magnitude for house numbers: absolute magnitude of the complex number and its logarithm; and (b) clock concepts for clock times: utterance referent (i.e., whole hour, half hour, or upcoming hour) and distance from referent (i.e., 0, 5, 10, or 15 min). We entered both magnitude and its logarithm into the regressions, because Brysbaert (1995) showed that log magnitude rather than absolute magnitude influences the time to process arabic integers ranging from 1 to 99.

Furthermore, the following groups of predictor variables, pertaining to morphophonological planning, were included for both house numbers and clock times: (a) numeral length: number of morphemes; number of syllables; number of phonemes; and (b) frequency: logarithm of the whole-form frequency, estimated from the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995) for the house numbers and from a Dutch newspaper-based small lexical database (TROUW Corpus) for the clock times. In addition, we used a second measure of frequency: logarithm of morpheme frequency (i.e., cumulative frequency of all the morphemes in the numeral, derived from the CELEX database; Baayen et al., 1995). Note that the factor of log morpheme frequency simultaneously incorporates the factor of numeral length, with longer numerals containing more morphemes and hence most probably having a higher log morpheme frequency. For example, the numeral 235, "tweehonderd*vijfen*dertig" ("two-hundred five and thirty"), has two morphemes more than the numeral 230, "tweehonderddertig" ("two-hundred thirty"), namely the morphemes "vijf" ("five") and "en" ("and"). Although not an ideal candidate, we believe that log morpheme frequency, along with the factor of whole-form frequency, provides us with some indication of whether or not the frequency of the complex numeral is playing a significant role in speech planning.

We fitted a multilevel multiple regression model (Pinheiro & Bates, 2000; see also Lorch & Myers, 1990) to the data with the logarithm[1] of naming latencies as dependent variable and participant as error stratum. In all analyses reported in this article, we first entered the total set of variables as predictors and assessed which made a major contribution. From there, we constructed the best fitting model.

For the house numbers, none of the above mentioned factors made a significant contribution in the regression analysis. As shown by the upper panel of Figure 1, naming latencies for the 12 number types were fairly constant, despite considerable differences in numeral length.

For the clock times, the best fitting regression model included the utterance referent (i.e., whole hour, half hour, or upcoming hour), distance from the referent (i.e., 0, 5, 10, or 15 minutes), and the logarithm of morpheme frequency as predictor variables. We observed significant effects for all three predictors. Naming latencies differed depending on the utterance referent, $\beta = 5.5744{\cdot}10^{-2}$, $t(4784) = 15.75$, $p < .0001$. As shown by the lower panel of Figure 1, utterances referring to the full hour (i.e., 00, 05, 10, 15) were produced much faster than utterances referring to the half hour (i.e., 20, 25, 30, 35, 40) and the coming hour (i.e., 45, 50, 55). Furthermore, naming latencies varied depending on the distance in minutes from the referent, $\beta = 1.1356{\cdot}10^{-2}$, $t(4784) = 2.96$, $p < .005$. Figure 1 shows that clock times showing a greater distance from the referent (e.g., 10 min) resulted in longer response latencies than those showing smaller distances from the referent (e.g., 5 min). A greater log morpheme frequency led to longer naming latencies, $\beta = 1.7784{\cdot}10^{-2}$, $t(4784) = 4.26$, $p < .0001$. All effects remained significant after partialling out the variance contributed by the other variables, $p < .01$ for all analyses. The standard deviation of the residual error in the model was 0.202. The correlation between the observed and the predicted naming latencies was 0.66, indicating a multiple R^2 of 44%. To summarize, the multiple regression results for the naming latencies suggest that conceptual preparation is needed for naming clock times presented in arabic digit format.

Gaze durations

Gaze durations were equivalent for both modes of response, $minF'(1, 19) < 1$. However, gaze durations depended on type, $minF'(11, 263) = 84.4$, $p < .01$. Furthermore, the effect of type varied with mode, $minF'(11, 279) = 33.27$, $p < .01$.

To gain further insight into the factors underlying the interaction between type and mode, we ran multiple regression analyses over the total data set. The same predictor variables as those used for the analysis of naming latencies were also entered in the multiple regression analysis of the gaze durations.

The best fitting regression model for the house number gaze durations included three predictor variables—namely, the logarithm of the whole-form frequency, the logarithm of the morpheme frequency, and numeral length (i.e., the number of morphemes and phonemes). We observed significant effects of all three predictors. A greater log whole-form frequency led to shorter gaze durations, $\beta = -1.0940{\cdot}10^{-2}$, $t(5398) = -4.59$, $p < .0001$. A greater log morpheme frequency likewise led to shorter gaze durations,

---

[1] In all multilevel multiple regression analyses reported in this paper, the logarithm of response latencies and gaze durations was taken as the dependent variable, in order to obtain a better fit, by reducing the variability caused by outliers in the data. Importantly, a similar pattern of results is obtained (i.e., best fitting models with an identical set of factors) when taking the response latencies and gaze durations as the dependent variable instead.

$\beta = -1.9173 \cdot 10^{-2}$, $t(5398) = -3.01$, $p < .005$. In contrast, longer numerals elicited longer gaze durations for the number of morphemes, $\beta = 8.5778 \cdot 10^{-2}$, $t(5398) = 13.84$, $p < .0001$, and for the number of phonemes, $\beta = 2.1064 \cdot 10^{-2}$, $t(5398) = 9.53$, $p < .0001$. All effects remained significant in sequential ANOVAs—that is, after partialling out the variance contributed by the other two variables, $p < .0001$ for all analyses. The standard deviation of the residual error in the model was 0.202. The correlation between the observed and predicted gaze durations was .77, indicating a multiple R^2 of 59%.

These multiple regression results suggest that when pairs of house numbers have to be named, gaze durations for the first house number in a pair are determined by morphophonological factors such as the whole-form and morpheme frequencies and numeral length (i.e., the number of phonemes and morphemes). As shown by the upper panel of Figure 1, gaze durations varied quite consistently with numeral length. House number utterances mentioning an additional morpheme (e.g., house numbers ending on a 5, like 235, requiring the response "tweehonderd*vijfen*dertig") were looked at much longer than shorter utterances (e.g., house numbers ending on a 0, like 230, requiring the response "tweehonderddertig"). These findings for gaze durations replicate earlier findings on the latencies of naming single complex numerals as house numbers (Meeuwissen et al., 2003). Also, they replicate the evidence obtained by Pynte (1974) for the effect of numeral length on gaze durations in the silent naming of two-digit numerals and the evidence on the effect of numeral length and frequency on gaze durations in the silent naming of one- and two-digit numerals obtained by Gielen et al. (1991).

The best fitting regression model for the clock time gaze durations included four predictor variables—namely utterance referent (i.e., whole hour, half hour, or upcoming hour), distance from the referent (i.e., 0, 5, 10, or 15 minutes), logarithm of whole-form frequency, and numeral length (i.e., the number of morphemes). We observed significant effects of all four predictors. Gaze durations differed depending on the utterance referent, $\beta = 1.17497 \cdot 10^{-1}$, $t(5487) = 24.93$, $p < .0001$. As shown by the lower panel of Figure 1, utterances referring to the full hour were looked at for a much shorter time than utterances referring to the half hour or the coming hour. Furthermore, gaze durations varied depending on the distance from the referent, $\beta = 2.4356 \cdot 10^{-2}$, $t(5487) = 6.58$, $p < .0001$. In Figure 1 it is shown that utterances mentioning a greater distance from the referent were looked at much longer than utterances mentioning a smaller distance. A greater log whole-form frequency led to shorter gaze durations, $\beta = -3.2600 \cdot 10^{-2}$, $t(5487) = -14.23$, $p < .0001$. Longer clock time utterances (i.e., in terms of number of morphemes) elicited longer gaze durations, $\beta = 1.02870 \cdot 10^{-1}$, $t(5487) = 12.20$, $p < .0001$. The standard deviation of the residual error in the model was 0.22. The correlation between the observed and the predicted gaze durations was .75, indicating a multiple R^2 of 56%. All effects remained significant after partialling out the variance contributed by the other three variables, $p < .0001$ for all analyses. To summarize, the multiple regression results showed that when pairs of clock times have to be named, the gaze durations for the first clock time in a pair are determined by both conceptual and morphophonological factors.

Discussion

The gaze durations but not the response latencies for the numeral pairs replicated the earlier findings for single complex numerals (Meeuwissen et al., 2003). The multiple regression results for the house numbers revealed a dissociation between naming latencies and gaze durations, with naming latencies being constant and gaze durations reflecting the influence of morphophonological variables. The naming latencies and gaze durations for the clock times both reflected the influence of conceptual and morphophonological variables. However, numeral length determined the gaze durations but not the response latencies for the clock times.

Thus, in both modes, dissociations were obtained between speech onset latencies (reflecting articulation) and gaze durations (reflecting planning), indicating that the participants planned and articulated simultaneously. However, the dissociation was greater for the house numbers than the clock times. Whereas none of the factors that determined gaze durations were reflected in the naming latencies for the house numbers, the dissociation between gaze durations and response latencies for the clock times concerned mainly numeral length. These results suggest that the extent of planning before initiating articulation is influenced by the nature of the utterance (house number vs. clock time).

EXPERIMENT 2

In the second experiment, the complex numeral pairs of Experiment 1 were presented in alphabetic format. Participants read the numerals aloud as house numbers or as clock times. For example, they said "op tweehonderd en driehonderdvijf" (literal translation: "at two hundred and three hundred five") in response to the pair TWEEHONDERD and DRIEHONDERDVIJF, and they said "om twee uur en vijf over drie" (literal translation: "at two o'clock and five past three") in response to the pair TWEE UUR and VIJF OVER DRIE. Gaze durations and reading latencies were measured for the first alphabetic numeral in a pair. We expected that the gaze durations would mirror the naming latency results found earlier (Meeuwissen et al., 2003), with only morphophonological factors determining the reading of house numbers and clock times. The first question was whether dissociations would be obtained between naming latencies and gaze durations in that gaze durations reflect the morphophonological factors even when reading latencies do not. Such dissociations would provide evidence for incrementality in reading complex numerals. The second question was whether the dissociations would depend on utterance mode—that is, whether the dissociations would occur for both the house numbers and the clock times.

Method

Participants

A total of 20 native speakers of Dutch, who did not participate in Experiment 1, took part in the experiment. They were undergraduate students of Nijmegen University. They were paid for their participation.

Materials, apparatus, design, and procedure

These were identical to those in Experiment 1, except that the pairs of numerals were now presented in alphabetic format for both response modes. By presenting numerals in alphabetic format, the perceptual input differed between the two response modes. Whereas, in Dutch, clock times are written with spaces between constituent words (e.g., VIJF VOOR HALF DRIE), house numbers are conventionally written as a connected string of letters (e.g., TWEEHONDERDVIJFENDERTIG).[2] As a result, this no longer allowed for a pair of house numbers to be presented horizontally on the computer screen. Therefore, we presented the alphabetic numerals in the upper and lower halves of the screen.

Analyses. A trial was considered invalid when it included a speech error, when a wrong oral response was produced, when the oral response was not fluent, or when the voice key was triggered incorrectly. Error trials were discarded from the analyses of the gaze durations and reading latencies. To analyse the speakers' gaze durations, we first classified their eye fixations as being located on the upper or lower alphabetic numeral or elsewhere. A fixation was counted as being

---

[2] It has been established that interword spaces benefit reading. However, a study by Inhoff, Radach, and Heller (2000), investigating the role of interword spaces in compound reading, has produced mixed results. They showed that interword spaces facilitate the access of the constituent word forms during the initial phase of compound reading, but hinder the subsequent specification of a conceptually unified compound meaning.

located on an alphabetic numeral if its coordinates lay within or on the outer contours of the numeral. On most trials, participants first fixated the upper and then the lower alphabetic numeral. Trials on which this was not the case were excluded from the analysis of the gaze durations and response latencies (this held for less than 1% of all trials).

Similar to Experiment 1, reading latencies, gaze durations, and errors were submitted to ANOVAs, with the crossed variables mode and type. Interactions of mode and type were further explored through multilevel multiple regressions on the total data set.

Results

Table 2 gives the mean reading latencies, their standard deviations, the first-pass gaze durations, their standard deviations, and the error percentages for the numeral types in the house number and clock time modes. Figure 2 displays the mean gaze durations and reading latencies for house numbers (upper panel) and clock times (lower panel).

Errors

Participants made an equal number of errors in both modes of response, $\mathit{minF}'(1, 36) < 1$.

Reading latencies

Reading latencies were equivalent for both modes of response, $\mathit{minF}'(1, 20) = 2.52$, *ns.* Furthermore, reading latencies did not depend on type, $\mathit{minF}'(11, 191) = 1.08$, *ns.* Furthermore, the effect of type did not depend on the mode, $\mathit{minF}'(11, 234) = 1.06$, *ns.*

Although the ANOVAs did not yield any significant results, for matters of consistency we carried out multiple regression analyses over the total data set. The same predictor variables as those in Experiment 1 were entered into the analyses. We fitted a multilevel multiple regression model to the data with the logarithm of reading latencies as dependent variable and participant as error stratum.

As in Experiment 1, none of the factors made a significant contribution to the reading latencies of the house numbers in the regression analysis. As shown by the upper panel of Figure 2, the

Table 2. *Mean latencies*[a] *for the vocal responses, gaze shifts*[a], *and error percentages, per mode and type for Experiment 2*

	Mode									
	House number					*Clock time*				
	Vocal		*Gaze*			*Vocal*		*Gaze*		
Type	*M*	*SD*	*M*	*SD*	*% Error*	*M*	*SD*	*M*	*SD*	*% Error*
00	531	150	609	208	4.8	491	127	519	153	4.2
05	543	152	886	282	4.9	505	151	843	250	6.5
10	539	159	897	340	7.3	516	139	797	220	3.8
15	522	156	956	270	5.4	515	144	788	201	4.6
20	541	159	953	294	4.9	529	149	989	255	5.0
25	535	163	1,253	335	7.0	530	157	1,008	254	6.9
30	536	156	939	282	6.1	495	131	555	184	4.4
35	540	160	1,221	304	6.0	515	153	1,069	289	4.8
40	533	152	1,009	329	7.1	525	161	1,019	250	2.9
45	538	161	1,266	315	6.3	513	143	778	208	5.8
50	529	153	974	332	6.0	517	149	768	224	4.6
55	534	158	1,303	337	8.5	520	140	780	200	7.1
All types	535	157	1,022	361	6.2	514	146	826	281	5.0

[a] In ms.

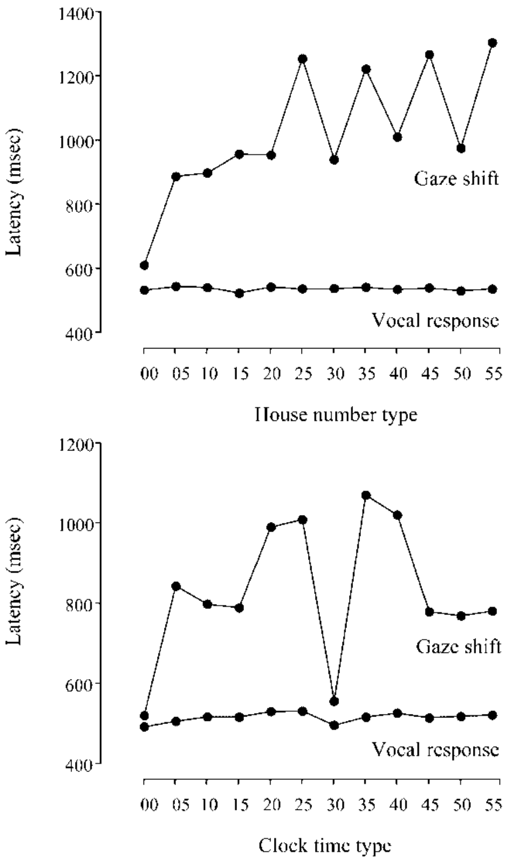

Figure 2. *Mean latencies of the vocal response and the shift of gaze in reading house numbers and clock times in Experiment 2.*

reading latencies did not differ much among the 12 number types, despite there being considerable differences in numeral length among types.

The same predictor variables as those in Experiment 1, including those coding the utterance referent and the distance from the referent in minutes, were entered into the analysis of the reading latencies of the clock times. We fitted a multilevel multiple regression model to the data with the logarithm of the reading latencies as dependent variable and participant as error stratum. The best fitting regression model for the clock time mode included one factor—namely, the logarithm of morpheme frequency. We observed significant effects for the predictor variable. A greater log morpheme frequency led to longer reading latencies, $\beta = 1.9728 \cdot 10^{-2}$, $t(4891) = 7.09, p < .0001$. The standard deviation of the residual error in the model was 0.171. The correlation between the observed and predicted reading latencies was .63, indicating a multiple R^2 of 40%. In contrast to Experiment 1, conceptual level factors such as utterance referent and distance from referent did not make a significant contribution, $F(1, 5168) = 0.46$, $p = .50$; $F(1, 5168) = 0.14, p = .71$, respectively. These multiple regression results suggest that morphophonological planning factors, but not conceptual factors, determine reading latencies for clock times.

Gaze durations

Gaze durations were on average 196 ms shorter for the clock times than for the house numbers, $minF'(1, 21) = 54.5$, $p < .01$, as indicated by Table 2. Moreover, gaze durations depended on type, $minF'(11, 211) = 124.04$, $p < .01$. Furthermore, the effect of type varied with response mode, $minF'(11, 180) = 45.57, p < .01$.

To gain further insight in the factors underlying the interaction between type and mode, we ran multiple regression analyses over the total data set. The predictor variables used for the analysis of the reading latencies were also entered in the analysis of the gaze durations.

The best fitting regression model for the house numbers included three predictor variables—namely, the logarithm of the whole-form frequency, the logarithm of the morpheme frequency, and numeral length (i.e., the number of morphemes, phonemes, and syllables). We observed significant effects for all three predictors in the model. A greater log whole-form frequency led to shorter gaze durations, $\beta = -3.4686 \cdot 10^{-2}$, $t(5550) = -15.99, p < .0001$. Likewise, a greater morpheme frequency led to shorter gaze durations, $\beta = -1.6246 \cdot 10^{-2}$, $t(5550) = -2.54$, $p < .05$. Longer numerals elicited longer gaze durations for the number of morphemes, $\beta = 5.2340 \cdot 10^{-2}$, $t(5550) = 7.20$, $p < .0001$, for the number of phonemes, $\beta = 1.3892 \cdot 10^{-2}$, $t(5550) = 4.43, p < .0001$, and for the number of syllables, $\beta = 6.2191 \cdot 10^{-2}$, $t(5550) = 7.79$, $p < .0001$. As shown by the upper panel of Figure 2, utterances mentioning an additional morpheme (e.g., house numbers ending on a 5, like 235, requiring the response “tweehonderd*vijfen*dertig”) were looked at for much longer than shorter utterances (e.g., house numbers ending on a 0, like 230, requiring the response “tweehonderddertig”). All effects remained significant in sequential ANOVAs—that is, after partialling out the variance contributed by the other two variables, $p < .0001$ for all analyses. The standard deviation of the residual error in the model was 0.178. The correlation between the observed and predicted gaze durations was .79, indicating a multiple R^2 of 63%.

To summarize, the multiple regression results show that in reading aloud pairs of house numbers gaze durations for the first house number in a pair are determined by morphophonological factors, such as the logarithm of whole-form frequency, the logarithm of morpheme frequency, and numeral length (i.e., the number of phonemes, morphemes, and syllables). However, conceptual-level factors, such as absolute magnitude and its logarithm, did not play a role.

The best fitting regression model for the clock times included three predictor variables—namely, the logarithm of the whole-form frequency, the logarithm of the morpheme frequency, and numeral length (i.e., the number of phonemes). We observed significant effects of all three predictor variables. A greater log morpheme frequency led to longer gaze durations, $\beta = 2.4197 \cdot 10^{-2}$, $t(5539) = 5.34$, $p < .0001$. In contrast, a greater log whole-form frequency elicited shorter gaze durations, $\beta = -2.3041 \cdot 10^{-2}$, $t(5539) = -10.37$, $p < .0001$. Furthermore, longer clock times (i.e., in terms of the number of phonemes) elicited longer gaze durations, $\beta = 5.5659 \cdot 10^{-2}$, $t(5539) = 43.70$, $p < .0001$. All effects remained significant after partialling out the variance contributed by the other two variables, $p < .0001$ for all analyses. The standard deviation of the residual error in the model was 0.186. The correlation between the observed and predicted gaze durations was .75, indicating a multiple R^2 of 56%.

To summarize, the multiple regression results show that when conceptual preparation is made unnecessary by presenting clock times in alphabetic format, gaze durations are determined by morphophonological factors only. Conceptual variables (utterance referent and distance in minutes) did not determine gaze durations for clock time reading.

Discussion

As in Experiment 1, we observed that the gaze durations for the house numbers were determined by morphophonological factors. Moreover, whereas the gaze durations for the clock times in Experiment 1 reflected both morphophonological and conceptual factors, the gaze durations in Experiment 2 only reflected morphophonological factors. This suggests that when clock times have to be read aloud, conceptual preparation is no longer required, which corroborates previous findings (Meeuwissen et al., 2003).

The absence of a conceptual involvement (i.e., utterance referent and distance in minutes) in the reading of clock times can be further illustrated by comparing the gaze durations for the full hour and the half hour between naming (Figure 1) and reading (Figure 2). As shown by the lower panel of Figure 1, the gaze durations were much shorter for the full hour than for the half hour (i.e., 633 ms vs. 920 ms, respectively). This supports the theoretical claim that conceptual transformations are carried out on the minute information (e.g., the 30 in 2:30 has to be transformed into “half to”) and the hour information (e.g., the 2 in 2:30 has to be transformed into “three”) for the half hour but not for the full hour (e.g., the 2 in 2:00 remains “two”). In contrast, no conceptual transformations are required in reading clock times (e.g., “half drie” has to be produced in response to HALF DRIE). Consequently, the gaze durations should only be determined by morphophonological factors, such as the number of morphemes to be planned for production. Since the full hours (e.g., DRIE UUR) and the half hours (e.g., HALF DRIE) are comparable in numeral length (e.g., they have the same number of morphemes), gaze durations should be comparable as well, which corresponds to what is empirically observed (see Figure 2).

Overall, clock times were looked at for a much shorter time than house numbers. This difference might be due to differences in input length (in terms of number of letters), as clock times were on average shorter in length (on average 13 letters) than house numbers (on average 19 letters). Furthermore, house numbers are conventionally written out as a connected string of letters, making them harder to discern.

Whereas the gaze durations for the numeral pairs replicated the earlier findings on the response latencies for single complex numerals (Meeuwissen et al., 2003), the response latencies did not. In both modes, dissociations were obtained between speech onset latencies (reflecting articulation) and gaze durations (reflecting planning), indicating that the participants planned and articulated simultaneously. As in Experiment 1, the dissociation was greater for the house numbers than for the clock times. Whereas none of the factors that determined gaze durations was reflected in the naming latencies for the house numbers, the dissociation between gaze durations and response latencies for the clock times concerned again mainly numeral length. These results suggest that the extent of structural frame planning before initiating articulation is influenced by the nature of the utterance (house number vs. clock time).

GENERAL DISCUSSION

We reported two eyetracking experiments that examined whether the extent to which individuals articulate and plan simultaneously is influenced by the production task (naming vs. reading) and the utterance mode (house number vs. clock time expression). Pairs of complex numerals had to be named (arabic format, Experiment 1) or read aloud (alphabetic format, Experiment 2) as house numbers and as clock times. In both tasks and modes, dissociations were obtained between speech onset latencies (reflecting articulation) and gaze durations (reflecting planning), indicating incrementality. Table 3 lists the factors that made a significant contribution to the multiple regression models for each task (naming, reading) and mode (house number, clock time). Whereas none of the factors that determined gaze durations was reflected in the latencies of reading (Experiment 1) and naming (Experiment 2) of the house numbers, the dissociation between gaze durations and response latencies for the clock times concerned mainly numeral length in both tasks. Thus, the dissociations were the same regardless of whether the utterance was already fully spelled out orthographically (Experiment 2) or not (Experiment 1). Moreover, the dissociations were

Table 3. *Summary of the multiple regression analyses*

Task	*Mode*	*Response latencies*	*Gaze durations*
Naming	House number	None	Log whole-form frequency Log morpheme frequency Numeral length
	Clock time	Utterance referent Distance in minutes Log morpheme frequency	Utterance referent Distance in minutes Log whole-form frequency Numeral length
Reading	House number	None	Log whole-form frequency Log morpheme frequency Numeral length
	Clock time	Log morpheme frequency	Log whole-form frequency Log morpheme frequency Numeral length

Note: For each task (naming, reading) and mode (house number, clock time), the factors are listed that made a significant contribution to the regression model.

the same regardless of whether the utterances required only morphophonological planning (the clock times in Experiment 2) or whether they also required conceptual-level planning (the clock times in Experiment 1). The dissociations were mainly affected by the nature of the utterance (house number vs. clock time).

The reason why the dissociation was greater for the house numbers than for the clock times may be that the structure of the house number utterances was more predictable than the structure of the clock time utterances. The structure of the house numbers was always x *honderd* z, whereas the structure of the clock times was x *uur*, y *over* x, y *voor* x, y *over half* x, y *voor half* x, *kwart over* x, and *kwart voor* x. The predictable structure of the house numbers allows for minimal planning before the onset of articulation, as the second term in the utterance ("honderd") is predetermined. When a speaker starts articulation after the number preceding the term "honderd" has been planned, there may be enough time to plan the rest of the house number during the articulation of the first number and the term "honderd". One would expect to obtain a fairly constant response latency pattern, because the particular numerals involved (i.e., "twee" to "negen") are comparable in length (e.g., in terms of number of morphemes and phonemes), as empirically observed.

Conceptual-planning factors, such as magnitude and its logarithm, did not determine gaze durations for naming and reading aloud house numbers. Instead, gaze durations for the first house number in a pair reflected the time to plan the corresponding word forms in both naming and reading aloud. The gaze durations for naming are in line with the evidence obtained by Pynte (1974) and Gielen et al. (1991) for silent naming. However, given the results of Brysbaert (1995), it is somewhat surprising that the magnitude of the complex numerals did not make a major contribution to the best fitting model for house number gaze durations in our experiments. The absence of a magnitude effect in our experiments cannot be due to a lack of statistical power, because we had 60 (20 participants × 3 repetitions) observations per number magnitude as compared with the 45 observations per number magnitude in the Brysbaert (1995) experiments. Moreover, we assessed the contribution not only of the absolute magnitude of the complex numeral, but also of its logarithm (according to Brysbaert, 1995, the reason why Gielen et al., 1991, did not find a magnitude effect was because they used magnitude instead of log magnitude). It is possible that we did not find a contribution of magnitude because of the high intercorrelations between morpheme frequency (which was a major predictor) and magnitude. Also, it is possible that an effect of magnitude was absent because of the task demands. Referring to the complex numerals as "house numbers" might have biased participants to access a different format of number representation (e.g., order rather than cardinality). Whatever the reason for the absence of an effect of magnitude, important for now is that the dissociation between production latencies and gaze durations mainly varied with the type of utterance (house number vs. clock time expressions).

To summarize, the reported experiments examined whether the extent to which individuals articulate and plan simultaneously is influenced by the production task (naming vs. reading) and the utterance mode (house number vs. clock time expression). In both tasks and modes, dissociations were obtained between speech onset latencies and gaze durations, indicating incrementality. Furthermore, whereas none of the factors that determined gaze durations were reflected in the latencies of reading and naming of the house numbers, the dissociation between gaze durations and response latencies for the clock times concerned mainly numeral length in both tasks. Thus, type of utterance (house number vs. clock time) mainly determined the degree of incrementality in speaking in the current experiments. These results highlight the importance of utterance structure in determining the degree of incrementality.

Returning to the broader issue of incrementality in speech production, we can conclude that the extent to which people speak incrementally (indicated by dissociations between speech onset latencies and gaze durations) is under strategic control (e.g., an incremental approach is more likely to be adopted by speakers when put under time pressure, Ferreira & Swets, 2002) and more likely to occur when long utterances have to be produced (Levelt & Meyer, 2000). In addition, from our own research, we can conclude that incrementality in speaking is more likely to occur when more than one object has to be named or read aloud (i.e., two complex numerals vs. a single complex numeral, cf. Meeuwissen et al., 2003) and when the structure of the utterance is more predictable (i.e., house numbers vs. clock times).

Original manuscript received 2 December 2003
Accepted revision received 12 February 2005

REFERENCES

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database* [CD-ROM]. Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Barrouillet, P., Camos, V., Perruchet, P., & Seron, X. (2004). ADAPT: A developmental, asemantic, and procedural model for transcoding from verbal to arabic numerals. *Psychological Review*, *111*, 368–394.

Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, *48*, 653–685.

Brysbaert, M. (1995). Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General*, *124*, 434–452.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335–359.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*, 57–84.

Gielen, I., Brysbaert, M., & Dhondt, A. (1991). The syllable-length effect in number processing is task-dependent. *Perception & Psychophysics*, *50*, 449–458.

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition*, *82*, B1–B14.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*, 274–279.

Inhoff, A. W., Radach, R., & Heller, D. (2000). Complex compounds in German: Interword spaces facilitate segmentation but hinder assignment of meaning. *Journal of Memory and Language*, *42*, 23–50.

Levelt, W. J. M., & Meyer, A. S. (2000). Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology*, *12*, 433–452.

Lorch, R. F., Jr., & Myers, J. L. (1990). Regression analyses of repeated measures data in cognition research. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 149–157.

Meeuwissen, M., Roelofs, A., & Levelt, W. J. M. (2003). Planning levels in naming and reading complex numerals. *Memory & Cognition*, *31*, 1238–1248.

Meeuwissen, M., Roelofs, A., & Levelt, W. J. M. (2004). Naming analog clocks conceptually facilitates naming digital clocks. *Brain and Language*, *90*, 434–440.

Meeuwissen, M., Roelofs, A., & Levelt, W. J. M. (2005). *Levels of speech planning in clock time production*. Manuscript submitted for publication.

Meyer, A. S., & Lethaus, F. (2004). The use of eye-tracking in studies of sentence generation. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

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, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, *66*, B25–B33.

Meyer, A. S., & Van der Meulen, F. F. (2000). Phonological priming of picture viewing and picture naming. *Psychological Bulletin & Review*, *7*, 314–319.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS (statistics and computing)*. New York: Springer.

Pynte, J. (1974). Readiness for pronunciation during the reading process. *Perception & Psychophysics*, *16*, 110–112.

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with "The language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language*, *41*, 416–426.