It's Time to Do the Math - Computation and Retrieval in phrase production.

Simone A. Sprenger[1] and Hedderik van Rijn[2]

[1] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

[2] Department of Experimental Psychology, University of Groningen, The Netherlands.

Accepted for publication in The Mental Lexicon

Address for correspondence:

Dr. S.A. Sprenger, University of Groningen, Faculty of Arts, Department of English Linguistics, Oude Kijk in 't Jatstraat 26 , 9712 EK Groningen, The Netherlands. Email: s.a.sprenger@rug.nl.

Acknowledgements:

The authors would like to thank R. Harald Baayen, Marc Brysbaert, Pim Levelt and Anastasios Sarampalis for their valuable comments on earlier versions of the manuscript.

Abstract

Research suggests that the lexicon stores high-frequent combinations of words (e.g., Arnon and Snider, 2010), thereby optimizing the balance between computation and retrieval during language comprehension. Here, we studied whether the production of multi-word expressions is optimized in a similar way. First, we measured speech onset latencies for Dutch clock-time expressions. Second, we developed a statistical model of these latencies, using two types of predictors: 1. the speech-onset latencies for arithmetic problems involved in Dutch clock time naming, and 2. the expressions’ Google frequencies. The resulting model explains 94% of the variance in our naming study. We conclude that phrase production shows the same frequency-driven balance between online computation and lexicon-related retrieval found in phrase comprehension.

Keywords: language production, retrieval versus computation, multi-word utterances, idiom processing, mental arithmetic

Language processing depends on the successful combination of memory retrieval and the application of rules, concerning both grammar and knowledge of the world. For example, if a friend asks us to meet her at the station at ten to three, we do not only need to recognize the words involved, but also need to combine them into a meaningful clock time, so that we can pick her up on time at 2:50 PM. For the ease and success with which we process this information, the particular processing unit involved is of crucial importance: do we analyze the phrase word-by-word, do we recognize the string as a realization of a particular construction, or do we retrieve the complete sequence as a familiar multi-word item? Research on language comprehension suggests that each of these levels of analysis can contribute to online processing. With respect to language production, the question about the size of the processing unit is of equal importance. Do we build each utterance from scratch, or do we take into account the fact that some combinations of words have a special meaning, or that they are simply used more often than others? Although the answers to these questions might seem trivial, finding empirical evidence to support them is not, as the restrictions that arise from quantitative experimental methods cannot easily be reconciled with spontaneous speech. Ideally, one would strive for tight control of the utterances without preventing speakers from planning them.

The clock-time-naming task (Bock, Irwin, Davidson & Levelt, 2003; Bock, Irwin & Davidson, 2004; Korvorst, Roelofs & Levelt, 2006, 2007; Meeuwissen, Roelofs & Levelt, 2003, 2004) seems to come close to this ideal. This variant of the picture-naming paradigm can be used to elicit complex spoken responses without additional training, due to the familiarity with clock faces and the task of telling the time. The resulting utterances vary with respect to formats and complexity, both within and between languages (Bock et al., 2003). But how much 'planning from scratch' is really involved in this task? After all, in modern societies the need to tell the time is ubiquitous, and so are the corresponding time expressions. They might therefore not be new in the true, generative sense. Instead, speakers may be sensitive to their frequencies and store them in long-term memory. At the same time, it seems plausible that at least some time expressions can involve a fair amount of processing (e.g., when an absolute time stimulus must be translated into a relative time expression). Therefore, we have chosen the clock-time naming task to study the difference between planning a phrase ‘from scratch’ and retrieving it from memory.

Bock et al. (2003) combined eye tracking and speech-onset measurements to explore the time course of clock-time planning and production. Replicating earlier findings in picture descriptions (e.g., Griffin & Bock, 2000; Meyer & Van der Meulen, 2000), they found that eye movements to areas that provide phrase content information are time-locked with the production of the corresponding phrases. That is, participants reliably moved their eyes from minute to hour regions, or vice versa, depending on the order of these elements in the corresponding expression. The observed patterns of eye movements hold across languages, clock-time formats, and type of display (analog or digital), suggesting that clock-time naming is subject to the same incremental formulation processes that have been observed in extemporaneous picture descriptions (Griffin & Bock, 2000). Thus, while offering a high level of control due to the constraints of the task itself, clock-time naming can be used to study natural complex-utterance production. Although the variety of different clock-time expressions is limited, the advantage lies in the possibility to study similar expressions of various levels of complexity within the same paradigm.

This can best be illustrated with the rules that apply to Dutch relative clock-time naming. Testing five-minute intervals, Bock et al. (2003) showed that native speakers of Dutch prefer a relative system that uses the full and the half hour as referents. Thus, Dutch speakers say *tien over h* ('ten past h') when the clock shows h:10 and *vijf voor half h+1* ('five before half h+1') when the clock shows h:25. According to Bock and colleagues, the referent changes at twenty[1] (h:20) to the half hour and at forty minutes (h:40) back to the (next) full hour, respectively. In addition, the reference hour that is named in the utterance (h or h+1) changes at twenty minutes past the full hour.

--- INSERT FIGURE 1 ---

Figure 1 shows how the different Dutch expressions relate to the clock times. It illustrates that the clock stimulus can elicit a range of verbal responses that differ with respect to their computational complexity. While some clock times can be read directly off the clock (e.g., 2:10 is *tien over twee*, 'ten past two'), others seem to require a considerable amount of computation (e.g., 6:23 is *zeven voor half zeven*, 'seven to half seven'). However, the set of possible utterance types is strictly limited and based on only a handful of possible phrase types. Accordingly, Bock et al. (2003) suggest that time expressions are a kind of non-figurative idiom, construction, or formula (such as defined by Kuiper, 1996), possibly represented in the form of superlemmas (i.e., lemmas that specify both the grammatical form and the words that belong to an idiom; Cieślicka, 2010; Kuiper, van Egmond, Kempen & Sprenger, 2007; Nordmann, Cleland

& Bull, 2010; Sprenger, 2003; Sprenger, Levelt & Kempen, 2006). In other words, they suggest that clock-time expressions have their own entries in the mental lexicon.

The Idiomatic Nature of Clock Times

Thinking of clock-time expressions in terms of idioms can indeed explain why deviations from the clock-time format are considered 'wrong', even if they are both semantically and syntactically correct (cf. Pawley & Syder, 1983). For example, the Dutch sentence *het is vijfendertig over drie* (‘it is thirty-five past three’) could in principle be a correct way of expressing 3:35. However, to a native-speaker of Dutch it sounds about as idiomatic as *he kicked the pail* (rather than *he kicked the bucket)* sounds to an American. He or she would say *het is vijf over half vier* (‘it is five past half four’), or, more rarely, *het is drie uur vijfendertig* (‘it is three o’clock thirty-five’) instead. However, there is no intrinsic reason for this distinction, other than that there are pre-specified formats in the speaker's long-term memory that the utterance must adhere to. The expression itself is arbitrary and unpredictable (cf. Keysar & Bly, 1999). This can be further illustrated by the differences between Dutch and German. The Dutch expression ‘ten to half seven’ (for 6:20) is non-idiomatic in the dialect of German that is spoken close to the Dutch border. There, the corresponding expression is ‘twenty past six’, using the current hour as referent. In contrast, 6:25 is expressed as ‘five to half seven’ on both sides of the border. Thus, although the two languages each use both referents (hour and half hour), the exact position of the turning point is idiosyncratic.

What does it mean if clock-time expressions are idioms? First of all, we would expect to find evidence for a processing advantage in comparison to sentences that have to be built from scratch. Although theories of idiom processing differ in their assumptions about the exact kind of information that is being stored and/or the exact level of representation, there is a general consensus in the psycholinguistic literature that they are stored as such in the mental lexicon. The finding that idioms are processed faster than comparable literal sentences – both in comprehension (e.g., Swinney & Cutler, 1979, Conklin & Schmitt, 2008) and in production (Cutting and Bock, 1997) – provides empirical support for this assumption.

In linguistics, the idea of direct idiom retrieval is linked tightly to approaches that emphasize the role of constructions (e.g., Fillmore, Kay & O’Connor, 1988; Goldberg, 1995; Jackendoff, 1995a, 1995b, 2002) or formulas (Kuiper, 1996). Jackendoff, for example, advocates a view in which the lexicon comprises all sorts of phrases that the speakers of a language are familiar with, including proper idioms, song titles, quotations, clichés, and the like. This rather broad definition of idioms (or rather, fixed expressions) can also be found in computational linguistics, where phrasal units are identified based on the frequency of lexical co-occurrence, rather than on the basis of their – often difficult to define – figurative character. Studies on the processing of these units demonstrate a similar processing advantage as in the case of classic idioms (see Shaoul & Westbury, 2011, for a review). For example, in a study that used self-paced reading and sentence recall, Tremblay, Derwing, Libben and Westbury (2011) showed that high-frequent lexical sequences (e.g., *in the middle of the*) are read faster and recalled more often than low-frequent sequences (e.g., *in the front of the*). The authors interpret these findings in favor of the storage and direct retrieval of these items.

From this general perspective it is conceivable that clock-time expressions are stored in the lexicon and that we should find the same processing advantage that has been observed for proper idioms and other high-frequent lexical sequences.

Clock-time naming as compositional process

Other work on clock-time naming has emphasized its compositional nature. While Bock et al. (2003) focused on what clock-time expressions have in common, Meeuwissen et al. (2003) explored the factors that account for reaction-time differences between individual time expressions in digital-clock-time naming. Similar to Bock et al. (2003), they studied standard five-minute intervals (standard times). Comparing relative clock-time-naming latencies (2:10, ‘ten past two’) to those of house-number reading (210, ‘two hundred and ten’), they concluded that the latencies for clock-time expressions depend on a combination of conceptual and morpho-phonological factors. Specifically, the predictor variables of a regression model for these latencies included the utterance referent, the distance from the referent, the logarithm of the whole-form frequency, and the logarithm of the morpheme frequency, all of which showed significant effects. This model explained 44% of the variance in standard time naming latencies. Based on their findings, Meeuwissen et al. propose a procedural semantics (Johnson-Laird, 1983) for Dutch relative clock-time naming, in which the referent is determined first, depending on the value of the minutes (mm) displayed. Separate procedures are proposed to determine the current hour (h) and the next hour (h+1), such that latencies are longer for minute values that refer to the next hour. Additional procedures carry out minute transformations for all cases in which the minute term cannot simply be read off the display.

Although the exact nature of these transformations has not been specified, Meeuwissen et al. (2003) provide us with a hypothesis about what happens in the early stages of clock-time processing: some kind of internal calculation is required before utterance formulation can start. At the very least, the minute value must be compared to

20 and –in seven out of twelve cases (given twelve different standard times) – the hour value must be increased by one.

Whereas the calculation of the hour terms seems rather straightforward, the proposed numerical transformations of the minute term can be achieved in either of two ways. One possibility is a pattern-matching mechanism that matches the required minute term to each stimulus (e.g., 5 is five, 25 is five, etc.). The other possibility is that speakers calculate the required terms using arithmetic procedures. In the latter case, the system can exploit the regularity of the relationship between stimulus and target term by storing one rule for each type of expression (i.e., those of the formats B, D, F, and H in Fig. 1). This option becomes especially attractive if we consider that the rules of relative clock-time naming do not only apply to the standard times: a single rule ("calculate minutes minus 30") could apply to all minutes between 19 and 29. In fact, the complete set of expressions can be represented by a limited number of clock time templates and their associated rules. Table 1 lists the corresponding formulas and the required calculations.

--- INSERT TABLE 1 ---

Taken together, the literature on clock-time naming provides us with two opposite hypotheses about the underlying planning processes, both of which have been derived from research on standard times. On the one hand, Bock et al. (2003) suggest that that clock-time expressions are idioms, which implies that they are retrieved from the mental lexicon in the form of pre-specified units. On the other hand, Meeuwissen et al.'s (2003) procedural grammar specifies that the minute term of a clock-time expression is the result of specific minute transformations, suggesting that clock-time expressions are compositional and computed on the spot (numerical transformation hypothesis).

In the following, we will argue that these opposite accounts are actually compatible with each other. In fact, they are both necessary to explain the behavioral differences between individual clock-time expressions, because the task of telling the time is accomplished by a flexible interplay of retrieval and computation. We support this claim with a combination of behavioral and corpus data. First, we present the results of a digital-clock-time naming study in which we measured the speech-onset latencies for all minutes of the hour. This expansion of the original clock-time-naming paradigm allows for contrasting the production of familiar (e.g., *vijf voor half zeven*, ‘five to half seven’) and less familiar time expressions (e.g., *drie voor half zeven*, ‘three to half seven’).

The resulting data set serves as the basis for a multilevel regression model in which we expand Meeuwissen et al.'s (2003) set of factors with the results of two additional studies. First, we measured speech-onset latencies in a mental-arithmetic task. The arithmetic problems were similar to those one would expect to be involved in clock-time naming, given the predictions of the numerical-transformation hypothesis. Using the arithmetic latencies to explain the variance in our first study, we put to the test the idea that telling the time involves arithmetic calculations. Second, we provide Google frequency counts of Dutch relative clock-time expressions in order to test our impression that some clock-time expressions are more common than others. Again, we use these data to explain the variance in clock-time naming latencies. If the time it takes to tell the time depends on the frequency of the clock-time expression, we take this as indication for a direct retrieval of this expression from the lexicon.

In other words, we assess the extent to which clock-time expressions can be considered fixed expressions or idioms by comparing the influences of arithmetic and frequency on speech onset latencies. After presenting these studies, we will discuss the theoretical implications of our results, with a special focus on the contributions of retrieval and computation to complex utterance production.

Study 1: Clock-Time-Naming Latencies for All Minutes of the Hour

The first study was devised to generate a new data set of clock- time naming latencies that would allow us to study the contribution of different processing strategies in more detail. Therefore, in contrast to the studies discussed above, we used an extended stimulus set that comprises all sixty possible time points in an hour.

Method

*Participants***.** Twenty-seven participants, all students of the University of Nijmegen, were tested. They were paid for their participation.

Materials and Design. The set of stimuli comprised all digital clock times from two o'clock (2:00) to nine fifty-nine (9:59), including the subset of standard times tested by Meeuwissen et al. (2003). The set was chosen such that all time points could be displayed with three-digit Arabic numerals ($h{:}m_1m_2$). The complete set of stimuli consisted of 480 unique items.

Procedure. Participants were tested individually. They were instructed to produce spoken clock times in response to a digital clock display on a computer monitor, using the relative clock-time format, and introducing each response with an initial om (at). Response latencies were measured by voice key.

At the beginning of each trial, a fixation cross was presented for 500 ms, followed by 150 ms blank screen. Then the clock-time stimulus was presented for 1000 ms. Speech-onset latencies were measured from clock presentation onset, with a deadline of 2500 ms. A new trial was initiated 1500 ms after voice-key triggering. Following a training session with 15 items, all stimuli were presented in random order in six individual blocks that were separated by a short pause. All utterances were checked for erroneous or missing responses and disfluencies. Data from two participants were removed from the data set, because the participants did not adhere to the instructions.

--- INSERT FIGURE 2 ---

Results and Discussion

The results of Study 1 are shown in Figure 2. The points connected by solid lines show the speech onset latencies per minute. The dashed line connects the standard times, indicating the effect one would expect for the minute terms if we had interpolated the non-standard times on the basis of the standard times. This dashed line replicates the pattern for standard times found by Meeuwissen et al. (2003), including an increase in reaction times as a function of the distance between a minute value and its referent. The complex pattern of reaction times for the combination of standard and non-standard times suggests that there are various different sources of variance in clock-time naming. Overall, reaction times seem to vary with utterance format, with longer reaction times for the more complex ones. However, the figure also shows that there is also a lot of variation *within* utterance formats, most notably the speedup for the standard times in comparison with the neighboring minute points, which has not been predicted by the set of factors included in Meeuwissen et al.'s model. We expect that expansion of the model with more specific assumptions about the type of calculations involved may help explaining these patterns.

Study 2: Arithmetic processing

In the second study, we follow-up on the possibility that clock-time naming involves mental arithmetic. To this aim, we studied arithmetic processing itself. As we discussed above, the numerical transformation hypothesis predicts that the language production system must interact with the general cognitive system in order to arrive at the correct minute and hour terms. In other words, speakers must calculate in order to tell the time.

Table 1 shows the four different types of minute calculation that are presumably involved in this process. Each of these functions relates to a different set of clock times and depends both on the referent (full or half hour), and the relative position to it (before or after). We assume that, like all cognitive operations, these calculations take a certain amount of time. The amount varies with the type of operation and with the overall complexity of the task, depending on the magnitude of the two operands, the magnitude of the solution, and the type of operation involved (for reviews on cognitive arithmetic, see for example Ashcraft, 1992, 1995; Campbell, 1995; Geary 1996). Calculation time can further be modulated by number-specific effects, such as parity (e.g., Dehaene, Bossini & Giraux, 1993). Most importantly however, it can easily be measured in terms of speech-onset latencies in a simple arithmetic task. With respect to relative clock time naming, our reasoning was as follows: if the minute term really results from arithmetic problem solving operations, calculation latencies should explain a significant portion of the variance in our clock-time-naming latencies. This prediction is supported by the findings of Ferreira & Swets (2002), who found that speech onset latencies for utterances in which subjects name the result of an arithmetic problem (e.g., *seven*, *seven is the answer*, or *the answer is seven*) vary with problem difficulty. We therefore measured calculation latencies in an independent study.

Method

All arithmetic problems associated with clock-time naming were constructed in four different formats. The design reflected the distribution of alternative formats for clock times in Dutch, as depicted in Figure 1, excluding the cardinal times (as producing these times does not involve arithmetic, e.g., *15* does not need to be calculated when the concept expressed is *a quarter past*). The time points *m* of the Formats *B, D, F*, and *H* were represented as $0 + m$, $30 - m$, $m - 30$, or $60 - m$, respectively. Participants were instructed to solve the arithmetic problem as fast as possible, introducing each verbal response with an initial *is* (as in *"is twenty-three" as response* to the stimulus "30 – 7 ="). There were 18 items each for the Formats B and H, and 9 items each for the Formats D and F. Each item was repeated four times. All stimuli were presented in random order in three individual blocks that were separated by a short pause. The experimental procedure was identical to the one applied in Study 1. Fifteen participants were tested.

--- INSERT FIGURE 3 ---

Results and Discussion

Average speech onset latencies measured in the arithmetic study are shown in Figure 3. The pattern shows that reaction times vary both between and within calculation types. Between calculation types, the main difference is that the addition problems $(0 + m =)$

are much faster than the subtraction problems in which one of the operands is *30 (30 - m =,* and *m - 30 =),* which are in turn faster than the subtraction problems that have *60* as an operand (*60 - m =).* Within calculation types, the pattern is relatively straightforward: a positive correlation between the difference of the two operands and the latency. However, this effect is reversed for the problems that result in the answer “five” or “ten”. That is, solving *40-30* or *30-20* is significantly faster than solving *35-30* or *30-25* ($F(2,20) = 43.34$, $p < 0.001$). Thus, for standard times the effect of mental arithmetic is diametrically opposed to the reaction time pattern that we observed both in Study 1 and in Meeuwissen et al.'s (2003) data. Below, we will show that the calculation times are nevertheless highly correlated with the speech onset latencies in Study 1.

However, before we turn to the regression model, we present a third study that was devised to attain an estimate of the frequencies of individual Dutch relative time expressions. Based on the assumption that high-frequent combinations of lexical items are represented in and retrieved from the mental lexicon in a single step, we hypothesize that clock-time-expression frequencies correlate with the corresponding speech onset latencies.

Study 3: Clock-time-expression frequencies

In western societies, time discipline is highly valued and clock-time expressions are ubiquitous. Thus, intuitively, the idea that clock-time expressions are idioms is not only supported by their idiosyncrasies, but also by the sheer frequency of use. It is conceivable that the probability with which an item becomes a fixed expression is directly related to its frequency (see also Sosa & MacFarlane, 2002; Sprenger, 2003; Wray, 2002). Once these representations have been formed, their activation level will be further modulated by frequency. Hence, given the above-mentioned processing advantage that arises from idiomaticity, the frequency of a given clock-time expression should be a good predictor for the corresponding production latency.

We therefore decided to assess the frequency of Dutch relative clock-time expressions by means of a Google search on Dutch web pages. The relationship between Google frequencies and the degree to which an expression is "fixed" can be illustrated with some familiar phrases. For example, the search query *nice to meet you* results in 18 million Google hits. Given its frequency, it stands a better chance to have a mental representation on its own than, for example, *delighted to meet you* (1.4 million Google hits) With respect to clock-time expressions, we expect to observe similar differences between standard and non-standard times, because schedules and appointments tend to be confined to the standard times (and the cardinal times in particular).

Method

We determined the frequencies of 103 clock time clusters of the format [x (alphabetic) [prep] [half]] (e.g., *vijf voor half*) by subjecting them to a Google search of Dutch-language web pages. All queries were submitted within 24 hours from the same IP address. The number of unique pages returned by Google was adjusted using the proportion of queries actually referring to clock times in a sample of 100 hits. If the number of pages was smaller than 100, all pages were inspected. We only counted unambiguous references to clock times in running text. That is, references to clock times in vocabulary lists, titles of pages, or computer programming code were omitted, whereas a sentence as *Het is vijf voor half zeven Nederlandse tijd* (“It is five to seven Dutch time”) in a travel journal (http://www.duppen.nl/fotografie/reisverhalen/canada/040530.html) did increase the count of *vijf voor half* by one. The range of queries covered all clock times in the formats A-D, as shown in Figure 1.

--- INSERT FIGURE 4 ---

Results

The results of the frequency count are shown in Figure 4. As expected, the cardinal times (marked with an additional dot) are more frequent than the other times. The remaining standard times are in turn more frequent than the non-standard times.

For a subset of time points, the figure shows two separate frequency estimates. These are time points that in Dutch are represented by two different time expressions each (the intervals from *h:16* to *h:25*, and from *h:35* to *h:44,* compare Figure 1). We decided to explicitly quantify the relative dominance of both formats. The resulting frequencies confirm the speakers' preferences, but also show that some of the non-preferred variants are frequent enough to be considered relevant competitors. Similar to the competition between words, this should translate into longer reaction times for the corresponding expressions (e.g., Roelofs, 1992, Levelt et al., 1999, Van Rijn & Anderson, 2003, Van Maanen & Van Rijn, 2007, Van Maanen, Van Rijn, & Borst, 2009).

Taken together, the observed pattern of frequencies confirms our expectations: the standard times, and in particular the cardinal times, are much more frequent than the remaining time points. This result fits well with our observation in Study 1 that these time points are named much faster than the others. However, a more detailed analysis is required to assess the influence of both frequency and arithmetic on clock-time naming.

Discussion

In our first study, we saw that if one considers the full range of Dutch relative clock-time expressions, the time it takes to tell the time varies more than Meeuwissen et al.'s (2003) procedural grammar predicts. Especially the difference between standard times and non-standard times shows that a comprehensive explanation of the observed latencies requires additional factors.

How can we explain that the standard times are so much faster than the rest? As we discussed in the introduction, the literature on clock-time naming offers two opposite hypotheses: Bock et al. (2003) consider clock-time expressions to be idioms and Meeuwissen et al. (2003) suggest that they require numerical transformations. Both accounts could in principle explain the observed differences. First, standard times might be named faster because they are used more frequently than the other time points and therefore have a representation on their own. In that case, they could be retrieved from memory as a whole, requiring neither minute term calculation nor syntactic formulation. If this assumed relationship between frequency and retrieval holds, the corpus frequencies of the clock-time expressions should be a reliable predictor of the associated naming latencies.

Second, it is possible that the standard times are named faster because the associated arithmetic operations (involving multiples of 5) are easier to perform than those that involve the other operands. This would fit into an account in which arithmetic operations are invoked by the language production system to calculate the minute term that belongs to a specific time point. If this account holds, the arithmetic latencies that we measured in our second study should be a reliable predictor for the clock naming latencies.

A third alternative however, is that clock-time expressions are a heterogeneous set of phrases that - depending on their frequency and therefore the strength of their representations - have different probabilities of direct retrieval. Therefore, in parallel with an attempt to retrieve them from the lexicon, syntactic formulation processes could build the expression according to the rules described above and trigger arithmetic procedures that calculate the required numeric terms.

Statistical model

In order to test these hypotheses, we performed a series of multilevel regression analyses (Baayen, Davidson & Bates, 2008) on the clock-time naming data collected in Study 1. We started with a simple model containing a number of morpho-syntactical and phonological predictors, and iteratively added additional factors. Each model was compared to the simpler model by a maximum likelihood-based χ^2 test. A table with all models can be found in the Appendix, but here we will focus on the final, preferred model. As this model was preferred by iterative χ^2 tests, all components contribute to the final fit. The details of this final model, which explains 94% of the between-item variance, are shown in Table 2.

--- INSERT TABLE 2 ---

The first two factors are not directly related to our hypotheses, but nevertheless can be expected to contribute to the variance in the speech latencies: the continuous factor *number of phonemes*[2] accounts for basic effects of phonology and the categorical factor *hour (2-9)* accounts for possible differences between hours.

The next group of factors accounts for differences in referent computation, following Meeuwissen et al. (2003). Taking the *current hour* reference as the baseline condition, we added *refHalf* (representing the half hour reference) and *refNext (*representing the next hour) as categorical factors. In addition, we included a continuous factor, inversely proportional to the distance (in minutes) between a time point and its referent, called *association*. As we discussed above, Meeuwissen et al. (2003) observed longer reaction times for a ten-minute distance to the referent than for a five-minute distance. In our own data set, we observe a similar rise of onset latencies with distance to referent. A possible explanation for this effect is the strength of the association between a minute and its referent, based on previous co-occurrences (e.g., Anderson & Lebiere, 1998).

In addition, we explored the contributions of mental arithmetic and frequency to the variance in speech latencies. *RT(calc)* represents the solution times in the arithmetic task. It has been scaled by subtracting the fastest average response time (to *0 + 5 =*) from that to the other problems. As arithmetic is only involved in non-cardinal time processing, the relative reaction time has been entered as an interaction with a binary cardinal-time factor.

The next factor addresses a structural difference between the arithmetic task and the clock-time-naming task. In the arithmetic task, the problems were presented in a ready-to-process form (e.g., *30 - 24 =)*. During clock-time naming however, we hypothesize that participants need to construct this format. This process is not sufficiently accounted for by the factors *refHalf* and *refNext*, because the assumed operations differ for time points before and past the half hour. The categorical factor *toHalf* represents the difference between the encoding of the subtraction problem for clock times before the half hour *(30-m)* and after the half hour *(m-30)*.

The next factor is related to the idiomaticity of clock-time expressions. To account for the ease with which a clock time expression can be retrieved from memory, we included the log-frequencies of the preferred and dispreferred formats as two separate factors. As can be seen in Table 2, the estimates for their parameters are in the expected direction. The negative estimate for the log frequency of the to-be produced clock times indicates that high-frequent expressions are produced faster than low frequent ones. Conversely, the positive estimate for the log frequency of the dispreferred formats confirms that clock times with higher frequent competitors are produced less fast than those with less frequent ones.

Finally, we addressed the relationship between the computational aspect of clock-time naming and the part that is played by storage and retrieval. From a processing point of view, calculating a minute term is not necessary when a clock-time expression can be retrieved as a whole from memory. Thus, the more frequent a time expression, the lower the probability that the speaker will need to finish the calculation, and the lower therefore the effect of referent and the calculation latencies. The negative value for the estimate of the interaction between the *Google-frequencies* and the *arithmetic* factor is in line with this processing view.

--- INSERT FIGURE 5 ---

Taken together, the full set of factors results in a relatively high proportion of explained variance and, as can be seen in Figure 5, captures all main effects in the empirical data. As removing any of the factors in this model results in a significantly decreased model fit, the model supports the hypothesis that the time it takes to produce Dutch relative clock-time expressions is the result of a joint effort of phrasal retrieval and computation.

General Discussion

The production of multi-word utterances is the result of the interplay of processes that retrieve and utilize facts and procedures from memory in general, and from the mental lexicon in particular. Clock-time naming is a task that allows studying these processes in more detail, albeit within a restricted domain.

The results of our first study show that the ease with which different types of clock times are produced differs widely, in a way that has not been predicted by Meeuwissen et al.'s (2003) procedural semantics. Our aim was to identify the sources of this variation and to show how the language production processes combine with general cognitive mechanisms when speakers produce multi-word utterances.

In particular, we were interested in the role of two opposing clock-time-naming strategies that have been discussed in the literature: Meeuwissen et al.'s (2003) numerical transformation hypothesis and Bock et al.'s (2003) idiom hypothesis. We identified two sets of independent predictors of the variance in our clock-time-naming data: mental arithmetic latencies for the problems that presumably arise during clock-time naming and the frequencies of Dutch relative clock-time expressions. A multi-level regression analysis of the naming data from Study 1showed that both factors substantially increase the amount of explained variance in clock-time-naming latencies.

From these findings, we can conclude that clock-time naming involves both computation of time expressions and the retrieval of clock time idioms, with frequency determining the relative weight of each strategy. The significant interaction between the arithmetic and frequency factors (i.e., the influence of the arithmetic factor is inversely proportional to the frequency of the expression) suggests that speakers follow a dual-route strategy in clock-time naming: high frequent expressions (like the standard times) are relatively easy to retrieve from the lexicon as a fully specified unit, while less frequent expressions more often need to be computed on the spot. For any given time expression, both routes may initially become active, but the probability of direct retrieval is a function of the expression’s frequency.

A dual-route account reconciles the numerical transformation hypothesis and the idiom hypothesis, be it with small changes to both. First, although Meeuwissen et al. (2003) correctly assumed the involvement of calculation in clock-time naming, our data suggest that their explanation is actually a more accurate description for the non-standard times than for the standard times that Meeuwissen et al. studied. Second, although a large proportion of all clock-time expressions is driven by retrieval-based idiomatic processes, producing a non-standard clock time involves arithmetic processing, and thus computation.

To our knowledge, the predicted relationship between the frequency of a fixed expression and the corresponding retrieval latencies during production has not been shown elsewhere. However, these results fit well into a pattern that has been observed for other domains of language processing, such as phrase comprehension and morphology. For example, Arnon & Snider (2010) recently reported frequency effects in the comprehension of compositional four-word phrases. In morphology, frequency effects for regular morphologically complex words have been taken as evidence for a parallel dual-route model of word processing. That is, regular full forms are simultaneously both retrieved from the lexicon and computed by a rule (Schreuder & Baayen, 1995; Baayen, Dijkstra & Schreuder, 1997; Baayen, McQueen, Dijkstra & Schreuder, 2003; Baayen, Schreuder, de Jong & Krott, 2002). Also, Tabak, Schreuder and Baayen (2010) report facilitatory frequency effects for the production of regular past-tense forms. But also for larger multi-word sequences, facilitating frequency effects have been observed. For example, Bannard and Matthews (2008) found an effect of word-sequence frequency on the speed and accuracy with which two to three-year old children are able to produce them in a repetition task. Taken together, these results suggest that proceduralization of the processing of phrases and words results in the storage of more complex forms that can directly be accessed in the lexicon. Our own data suggest that digital clock time processing proceeds along similar lines.

What could the two routes in clock-time naming look like? If we assume that high frequent clock-time expressions are represented as idioms or *fixed expressions*, the first, direct route requires the activation of a specific clock time concept and the retrieval of the corresponding lexical entry. This entry should specify the minute term, the referent, and possibly the syntactic structure that connects them. Such an account fits well with our observation that the production latencies for the standard times are mainly dependent on the frequency of the complete expression.

With respect to the hour terms, our speech onset latencies are not well suited to study the way in which they are processed. Following Bock et al.'s (2003) suggestion that clock-time naming involves hierarchical incrementality, preparation for the minute term might subsume partial preparation for the hour term. Also, the required calculations are relatively simple *(h' is either h or h+1).* Therefore, we suggest representing it with an open slot that can be filled while the first part of the utterance is being encoded. If we assume in addition that filling this slot is optional and context dependent, this could explain the occurrence of “incomplete” constructions such as *tien voor half* (‘ten to half’) or *vijf voor* (‘five to’) that Dutch speakers use when the hour is common ground or irrelevant (e.g., *de bus vertrekt altijd om tien voor half*, ‘the bus leaves always at ten to half’).

The second route, which is responsible for those times that have to be constructed online, can be implemented in different ways. Any solution however has to fulfill two constraints: First, the resulting expression must follow the rules of Dutch relative clock-time naming. Second, the processes involved must trigger arithmetic processing in the general cognitive system and integrate the results into the structure produced. The question is whether these constraints are implemented by the general cognitive system or by the language processing system itself. Is the way in which we tell the time rooted in how we perceive and process time itself, or is it just the way in which we happen to express it? That is, do speakers of Dutch *conceptualize* a given time in terms of minutes relative to the half hour or hour, or is it the linguistic frame that requires them to put it that way?

In the first case, clock time formulation follows the standard path of syntactic formulation. That is, it proceeds from a preverbal message to the activation of separate lexical entries and the online generation of a syntactic structure. The specific form of the expression results from the way in which speakers conceptualize a given time stimulus. Thus, speakers must encode the relation of the minutes and hours to each other and apply the rules of Dutch clock-time naming before generating the preverbal message. For example, they must identify the referent, decide in which way the minute relates to it (e.g., before or after the half hour), and then calculate the minute term on the basis of that input, using language-independent arithmetic procedures (Brysbaert, Fias & Noël, 1998). Subsequently, speakers can activate the corresponding lexical concepts. A drawback of this explanation is that it cannot easily explain why for example *zes minuten voor drie* (‘six minutes to three’, with the minute unit made explicit) is grammatical, but **zes voor drie uur (*‘six to three hour’, six to three o'clock) is not. In other words, there are structural and productive constraints to clock-time expressions that are difficult to explain on the level of conceptualization alone.

Alternatively, one can locate the intricacies of non-standard time naming within the linguistic processing system. For example, following the construction grammar approach (e.g., Fillmore, Kay and O'Connor, 1988), we can describe non-standard times in terms of an abstract idiomatic construction that is stored as such in the lexicon. Thus, similar to, for example, the *the X-er the Y-er* construction *(e.g., the bigger they come, the harder they fall)* or the *let alone* construction *(e.g., Max won't eat shrimp, let alone squid; I won't touch, let alone buy meat)*, the lexicon may contain a *minute relative to hour* construction of the form *[[m'] [voor, over] [half] [h']].* This construction has open slots for the minute term, the preposition, the half hour term, and the hour term. In this scenario, speakers first need to retrieve the construction and subsequently set the variables (*voor* vs. *over*, half or full hour) and fill the open slots.

In contrast to the first solution, this approach not only acknowledges the linguistic nature of the constraints that apply to non-standard time naming, but also captures the close relation between standard and non-standard times: following the terminology of Fillmore et al. (1988), we can think of the abstract construction for the low frequent time expressions as a *formal idiom* that serves as host to the *substantive* or *lexically filled idioms* that have been specified by virtue of their frequency.

On the downside, this second solution bears the cost of additional storage. However, as we have seen in the examples from morphology and phrase comprehension, storage capacity does not seem to be a major limiting factor in language processing and even storage of regular constructions can be a sign of efficiency if it helps to speed up processing (see also Nooteboom, Weerman & Wijnen, 2002). As Libben (2005, p.271) put it in the context of morphology, the lexicon may be said to be *hungry* in the sense that it has been “designed to make the greatest number of potentially useful representations and analyses available to other components of the cognitive system”.

At present, our data cannot distinguish between the two approaches to the construction of clock-time expressions. However, both are compatible with our observation that less frequent time expressions rely heavily on computation: in addition to conceptualization and/or syntactic processing, both solutions require the speaker to calculate. In contrast, the standard times can readily be retrieved from the lexicon[3].

Taken together, our results show that, similar to other levels of language processing, the production of complex spoken utterances is brought about by a highly flexible processing system in which the balance of retrieval and computation has been optimized.

References

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62,* 67–82.

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*, *44*(1–2), 75–106.

Ashcraft, M. H. (1995). Cognitive psychology and simple arithmetic: A review and summary of new directions. *Mathematical Cognition*, *1*(1), 3–34.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412.

Baayen, R.H., Dijkstra, A., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language, 37,* 94-117.

Baayen, R. H., McQueen, J., Dijkstra, T. and Schreuder, R. (2003). Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In Baayen, R. H. and Schreuder, R. (Eds.), *Morphological structure in language processing* (pp. 355-390), Berlin: Mouton.

Baayen, R. H., Schreuder, R., De Jong, N. H., and Krott, A. (2002). Dutch inflection: the rules that prove the exception. In Nooteboom, S., Weerman, F. and Wijnen, F. (Eds.), *Storage and Computation in the Language Faculty* (pp.

61-92), Dordrecht: Kluwer.

Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, *48*(4), 653–685.

Bock, K., Irwin, D. E., & Davidson, D. J., (2004). Putting first things first. In J. M. Henderson & F. Ferreira (Eds.): *The interface of language, vision and action. Eye movements and the visual world* (pp. 249-277). New York: Psychology Press.

Brysbaert, M., Fias, W., & Noël, M.-P. (1998). The Whorfian hypothesis and numerical cognition: is ‘twenty-four’ processed in the same way as ‘four-and-twenty’? *Cognition, 66*, 51-77.

Campbell, J. I. D. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition*, *1*(2), 121–164.

Cieślicka, A. (2010). Formulaic Language in L2. Storage, retrieval and production of idioms by second language learners. In M. Pütz and L. Sicola (Eds.), *Cognitive processing in second language acquisition: inside the learner’s mind* (pp. 149-168). Amsterdam: John Benjamins Publishing Co.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics 29*(1), 72–89.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson’s method. *Tutorial in Quantitative Methods for*

Psychology, 1(1), 42–45.

Cutting, J., & Bock, K. (1997). That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory and Cognition*, *25*, 57–71.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*(3), 371–396.

Dell, G. S. (1986). A spreading activation theory of retrieval in language production. *Psychological Review*, *93*, 283–321.

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*(1), 57-84.

Fillmore, C. J., Kay, P., and O'Connor, M.K. (1988). Regularity and idiomaticity in grammatical constructions: the case of let alone. *Language, 64,* 501–538.

Geary, D. C. (1996). The problem-size effect in mental addition: Developmental and cross-national trends. *Mathematical Cognition*, *2*(1), 63–94.

Goldberg A.E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*(4), 274–279.

Jackendoff, R. (1995a). The Boundaries of the Lexicon. In M. Everaert, E.-J. van der Linden, A. Schenk, and R. Schreuder (Eds.). *Idioms: Structural and Psychological Perspectives* (pp. 133-165). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jackendoff, R. (1995b). The Architecture of the Language Faculty. Boston: MIT Press.

Jackendoff, R. (2002). What’s in the Lexicon? In S. Nooteboom, F. Weerman, and F. Wijnen, (Eds.), *Storage and Computation in the Language Faculty* (pp. 23-58). Dordrecht: Kluwer.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge: Cambridge University Press.

Keysar, B., & Bly, B. (1999). Swimming against the current: Do idioms reflect conceptual structure? *Journal of Pragmatics, 31,* 1559-1578.

Korvorst, M., Roelofs, A., & Levelt, W. (2006). Incrementality in naming and reading complex numerals: Evidence from eyetracking. *The Quarterly Journal of Experimental Psychology, 59*(2), 296-311.

Korvorst, M., Roelofs, A., & Levelt, W. (2007). Telling time from analog and digital clocks: a multiple-route account. *Experimental Psychology*, *54*(3), 187–91.

Kuiper, K. (1996). *Smooth talkers*. Mahwah, NJ: Lawrence Erlbaum Associates Erlbaum Associates Mahwah, NJ.

Kuiper, K., Van Egmond, M. E., Kempen, G., & Sprenger, S. A. (2007). Slipping on superlemmas: Multi-word lexical items in speech production. *The Mental Lexicon*, *2,* (313-357).

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(01), 1–38.

Libben, G. (2005). Everything is Psycholinguistics: Material and methodological considerations in the study of compound processing. *The Canadian Journal of Linguistics, 50*(1/2/3/4), 267-283.

Meeuwissen, M., Roelofs, A., & Levelt, W. J. M. (2003). Planning levels in naming and reading complex numerals. *Memory & Cognition*, *31*(8), 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.

Meyer, A. S. & Van der Meulen, F. (2000). Phonological priming effects on speech onset latencies and viewing times in object naming. *Psychonomic Bulletin & Review, 7*(2), 314-319.

Nooteboom, S. G., Weerman, F. P., & Wijnen, F. N. K. (Eds.) (2002). *Storage and Computation in the Language Faculty*. Dordrecht: Kluwer. Studies in theoretical psycholinguistics; 30.

Nordmann, E., Cleland, A.A., & Bull, R. (2010). The cat's got your tongue: Using the tip-of-the-tongue state to investigate the 'superlemma' model of idiom production. *Poster presented at the 6th International Workshop on Language*

Production, Edinburgh, Scotland.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–225). London: Longman.

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

Schreuder, R. & Baayen, R.H. (1995). How complex simplex words can be. *Journal of Memory and Language, 37,* 118-139.

Shaoul, C. & Westbury, C. (2011). Formulaic Sequences. Do they exist and do they matter? *The Mental Lexicon, 6*(1), 171-196.

Sosa, A. V. & MacFarlane, J. (2002) Evidence for frequency-based constituents in the mental lexicon: collocations involving the word of. *Brain and Language, 83*(2), 227-236.

Sprenger, S. A. (2003). *Fixed expressions and the production of idioms* (Doctoral dissertation). Retrieved from http://hdl.handle.net/2066/63832.

Sprenger, S. A., Levelt, W. J. M., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language*, *54*(2), 161–184.

Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning & Verbal Behavior, 18*, 523-534.

Tabak, W, Schreuder, R. & Baayen, R.H. (2010). Producing inflected verbs: A picture naming study. *The Mental Lexicon, 5*(1), 22-46.

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning, 61*(2), 569-613.

Van Maanen, L., & Van Rijn, H. (2007). An accumulator model of semantic interference. *Cognitive Systems Research*, *8*(3), 174–181.

Van Maanen, L., Van Rijn, H., & Borst, J.P. (2009). Stroop and picture-word interference are two sides of the same coin. *Psychonomic Bulletin & Review, 16*(6), 987-999.

Van Rijn, H., & Anderson, J. R. (2003). Modeling lexical decision as ordinary retrieval. In F. Detje, D. Doerner, & H. Schaub (Eds.), *Proceedings of the fifth international conference on cognitive modeling* (pp. 207–212). Bamberg: Universitätsverlag Bamberg.

Wray, A. (2002). *Formulaic language*. Cambridge: Cambridge University Press.

Footnotes

[1] Note that speakers may occasionally change referents at :16 and at :39 (See areas "B or D" and "H or F" in Figure 1.)

[2] Note that originally, we also included the *number of morphemes* and *log morpheme frequency* (following Meeuwissen et al., 2003). However, model comparisons showed that these factors do not contribute sufficiently to the model fit to warrant inclusion in the final, preferred model. See the Appendix for the respective model comparisons.

[3] A computational model implementing this idea can be requested from the second author.

Tables

Table 1. Formulas in Dutch relative clock-time naming. Stimulus: h:mm, e.g. 5:23. m' = minute term, h' = hour term. English translations: uur = o'clock; kwart = quarter; over = past; and voor = to; prep=preposition.

Range	Formula	Calculation	Example
h:00	*h* uur	–	1:00 ⇒1 uur
h:15, h:45	kwart *prep h'*	if mm==15 ⇒ *prep*=over; *h'*=h if mm==45 ⇒ *prep*=voor; *h'*=h+1	2:15 ⇒ kwart over 2 4:45 ⇒ kwart voor 5
h:30	half *h'*	h'=h+1	3:30 ⇒ half 4
h:00 .. *h:19* *h:41* .. *h:59*	*m' prep h'*	if mm < 20 ⇒ *m'*=mm; *prep*=over; *h'*=h if mm > 40 ⇒ *m'*=60-mm; *h'*=h+1; *prep*=voor	2:07 ⇒ 7 over 2 7:51 ⇒ 9 voor 8
h:20 .. *h:29* *h:31* .. *h:40*	*m' prep* half *h'*	if (mm > 19 and mm < 30) ⇒ *m'* = 30-mm; *h'*=h; *prep*=voor if (mm > 30 and mm < 40) ⇒ *m'* = mm-30; *h'*=h+1; *prep*=over	3:20 ⇒ 10 voor half 4 8:37 ⇒ 7 over half 9

Table 2. Overview of the estimates *(β)*, the upper and lower 95% Bayesian highest posterior density (HPD) confidence intervals, and *p*-values based on the MCMC posterior distribution and on the *t*-distribution (determined using pvals.fnc with 10000 samples, Baayen, Davidson, Bates, 2008) of the fixed factors entered in linear mixed-effect models. Apart from these fixed effects, the model contained random effects reflecting by-subject adjustments to the intercept.

Predictors	β	HPD 95%		p
		lower	upper	MCMC-based
Intercept	6.35	6.2940	6.4064	0.0001
Number of phonemes	0.028	0.0222	0.0330	0.0001
Hour	*			
refHalf	-0.011	-0.0265	0.0023	0.1038
refNext	0.121	0.0993	0.1446	0.0001
Association	0.004	0.0025	0.0056	0.0001
RT(calc, noncardinal)	0.0004	0.0004	0.0005	0.0001
toHalf	0.026	0.0116	0.0408	0.0006
log(Frequency)	-0.0005	-0.0042	0.0030	0.8172
log(Competing Freq)	0.014	0.0087	0.0183	0.0001
RT(calc, noncardinal) x log(Freq)	-0.0005	-0.0001	-0.0000	0.0001
Model Evaluation:				
Individual RTs R^2	0.41			
Averaged RTs R^2	0.94			

Table 3. Statistical model estimations, fit and comparisons. All effects are significant at the $p < 0.001$ level, except effects marked with *, which are significant at the $p < 0.05$ level. No estimations are presented for hour, as each hour [2:9] has a separate estimated β and the individual values have no theoretical relevance. The averaged RTs R^2 is the explained variance when for both empirical data and predictions a correlation is calculated over the average latencies per minute. The χ^2 values reflect the comparison between model n and n-1.

Model:	1	2	3	4	5
Predictors	β	β	β	β	β
Intercept	5.74	6.13	6.29	6.44	6.35
Number of phonemes	0.036	0.034	0.035	0.037	0.028
Number of morphemes	0.074	0.007	-0.008	-0.030	
log(Morpheme Frequency)	0.043	0.015	0.005	0.0002	
Hour	*	*	*	*	*
refHalf		0.017	0.026	0.015	-0.011
refNext		0.190	0.096	0.109	0.121
Association		0.007	0.004	0.004	0.004
RT(calc, noncardinal)			0.0004	0.0003	0.0004
toHalf			0.023	0.030	0.026
log(Frequency)				-0.010	-0.0005
log(Competing Freq)				0.016	0.014
RT(calc, nc) x log(Freq)					-0.0005
Model Evaluation:					
Individual RTs R^2	0.34	0.40	0.41	0.41	0.41
Averaged RTs R^2	0.58	0.88	0.93	0.94	0.94
Model Comparisons:					
Df	13	16	18	20	19
logLik	189.28	713.77	798.44	818.48	829.59
χ^2		1049.0	169.34	40.07	

Figure Captions

Figure 1. Relative Dutch clock-time expressions and their distribution across the hour. *h* and *m* represent the actual hour and minute values.

Figure 2. Average speech onset latencies and associated corrected standard errors (Cousineau, 2005) for relative clock-time naming in Dutch (Study 1). The dashed line connects the standard times (represented by a larger dot) tested by Meeuwissen et al. (2003).

Figure 3. Average speech onset latencies and corrected standard errors (Cousineau, 2005) for the answers to simple arithmetic problems that have been derived from the rules of relative Dutch clock-time naming.

Figure 4. *Google* frequencies of Dutch relative clock-time expressions. The solid line represents frequencies of expressions as elicited from the participants in Study 1 (e.g., *elf over half X*, eleven over half X); the dashed line represents alternative expressions for the same minute term (e.g., *negentien voor X*, nineteen before X). The x-axis denotes the minute term and the y-axis denotes log frequency.

Figure 5. Empirical data from Figure 2 and fit of the statistical model presented in Table 2.

Figures

Figure 1

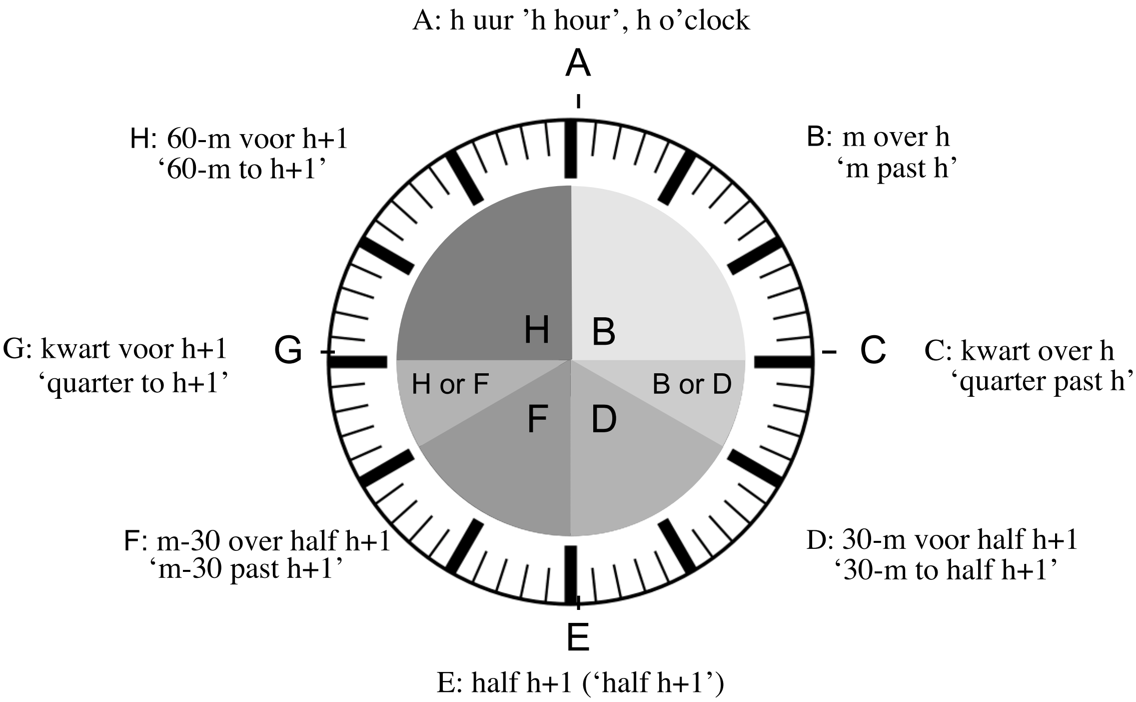


Figure 2

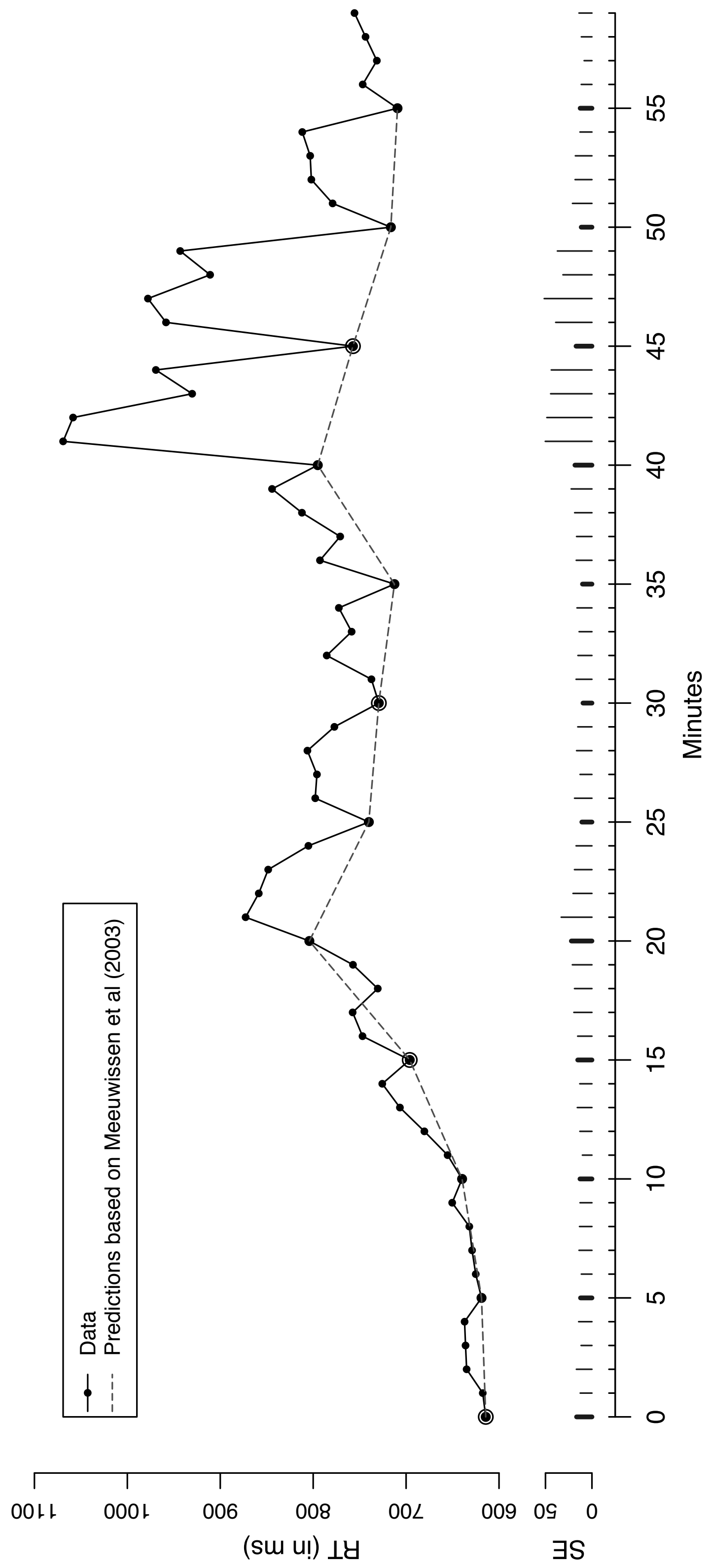

Data
Predictions based on Meeuwissen et al (2003)
Minutes
0
5
10
15
20
25
30
35
40
45
50
55
RT (in ms)
600
700
800
900
1000
1100
SE
0
50

Figure 3

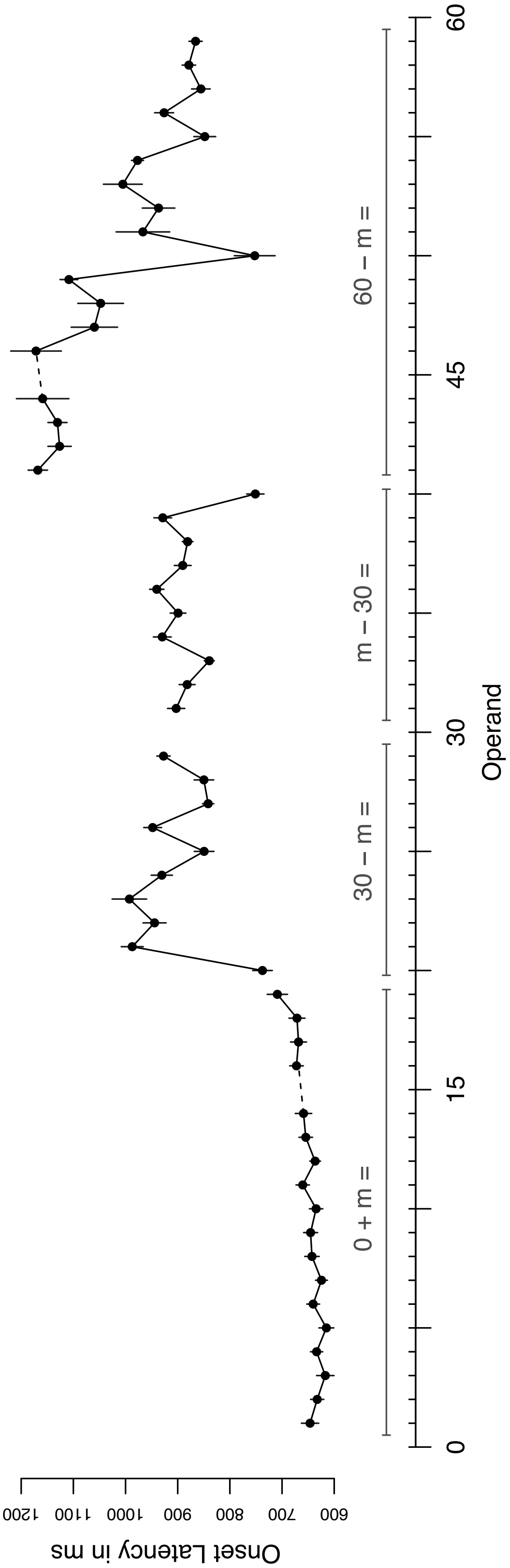

Onset Latency in ms
600
700
800
900
1000
1100
1200
0 + m =
30 − m =
m − 30 =
60 − m =
0
15
30
45
60
Operand

Figure 4

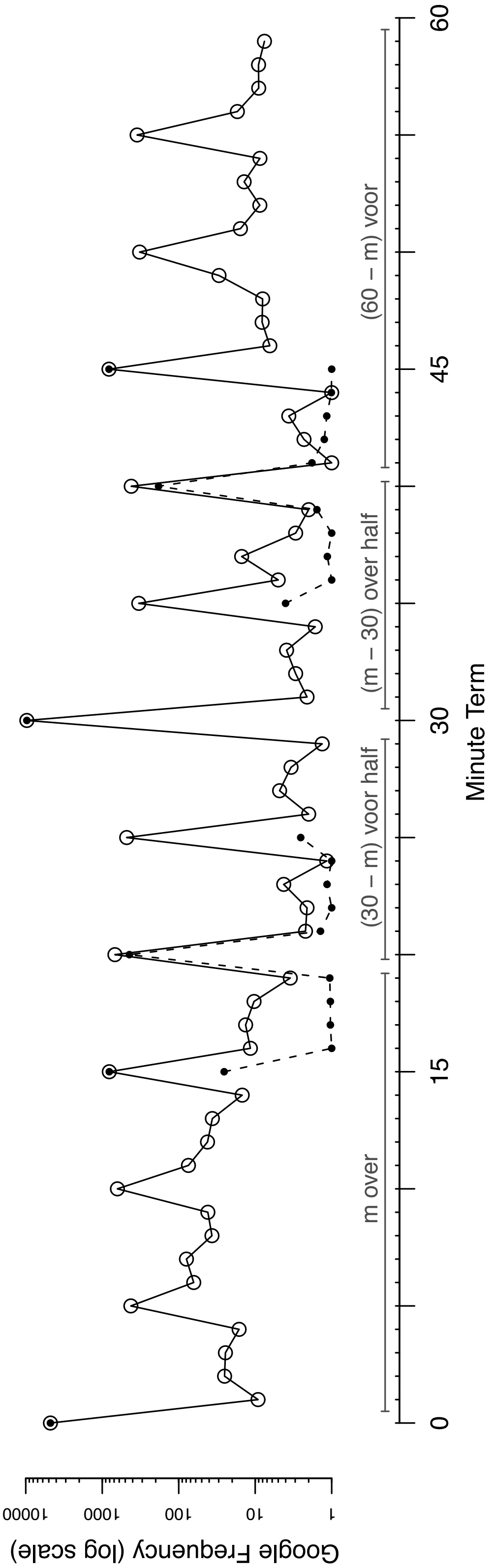
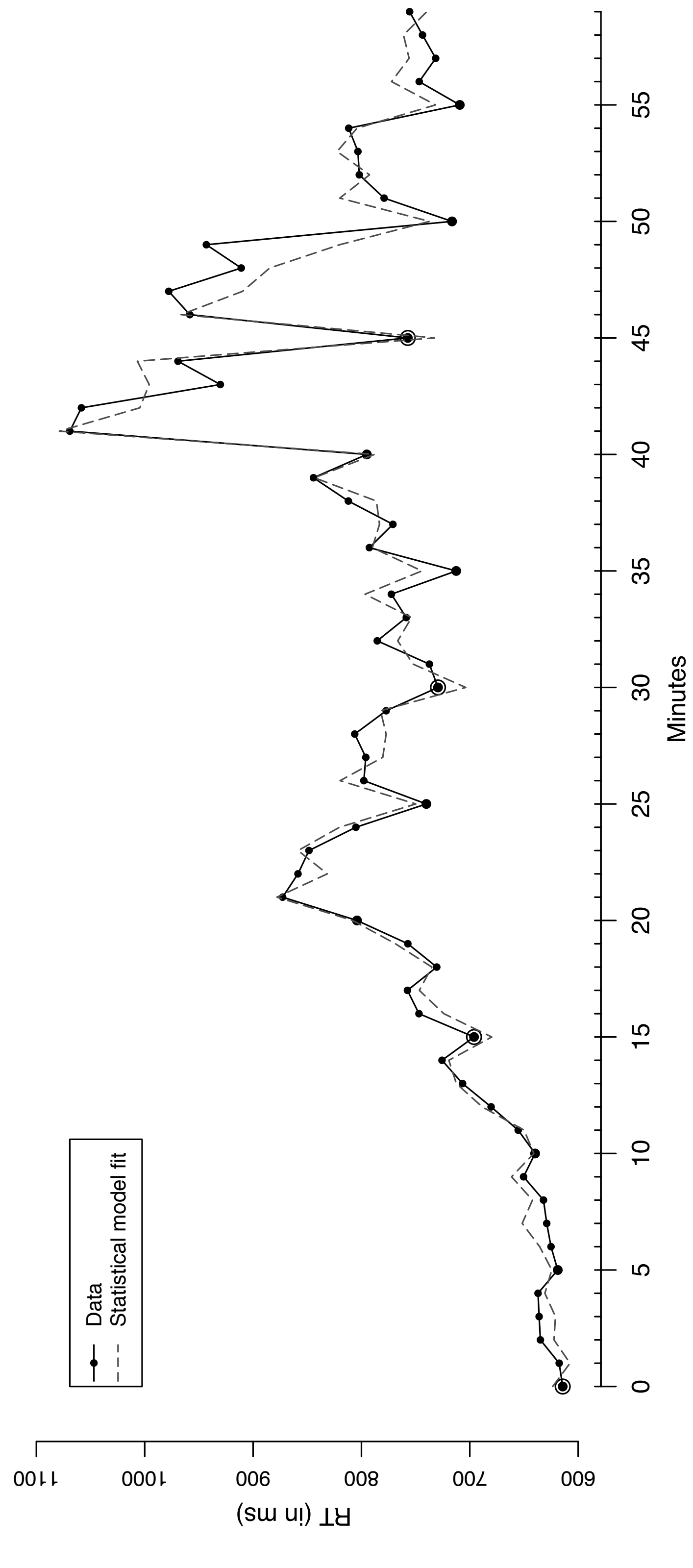

Google Frequency (log scale)
1
10
100
1000
10000
Minute Term
0
15
30
45
60
m over
(30 – m) voor half
(m – 30) over half
(60 – m) voor

Figure 5


Data
Statistical model fit
RT (in ms)
600
700
800
900
1000
1100
Minutes
0
5
10
15
20
25
30
35
40
45
50
55