Title
Clock Time Naming: Complexities of a Simple Task

Permalink
https://escholarship.org/uc/item/1tm664b0

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN
1069-7977

Authors
Sprenger, Simone A.
van Rijn, Hedderik

Publication Date
2005

Peer reviewed

Clock Time Naming: Complexities of a Simple Task

Simone A. Sprenger (simone.sprenger@mpi.nl)
Max Planck Institute for Psycholinguistics; P.O. Box 310
NL-6525 AH Nijmegen, The Netherlands

Hedderik van Rijn (hedderik@ai.rug.nl)
University of Groningen, Department of Artificial Intelligence; Grote Kruisstraat 2/1
NL-9712 TS Groningen, The Netherlands

Abstract

Performance in relative clock time naming (e.g., pronouncing 3:50 as "ten to four") has been described as depending on three factors: reference hour determination, minute transformation, and an additional distance component (Meeuwissen, Roelofs & Levelt, 2003). However, this model does not specify the cognitive operations that are responsible for the distance effect. We present three hypotheses about the factors that determine clock time latencies: physical distance, arithmetics, and frequency of the expression. Three experiments and a corpus analysis that test these hypotheses are presented. Regression models of speech onset latencies for an extended set of clock times show clear contributions of all three factors and explain most of the variance associated with this task.

Keywords: clock time naming, language production, phrases.

Clock time naming

In a series of recent papers (Bock, Irwin, Davidson & Levelt, 2003; Meeuwissen, Roelofs & Levelt, 2003, 2004) clock time naming has been established as a new paradigm to study the production of multi-word utterances. It can be seen as a special variant of the picture naming paradigm that allows to elicit complex spoken responses without additional training, due to the universal familiarity of clock faces and the task of telling the time.

Bock et al. (2003) showed that while the stimulus is universal, the format of the utterance that is used to express time varies both between and within languages. Most importantly, languages like English and Dutch (among many others) offer two alternative ways of telling the time: either using an absolute system (*three fifteen*) or using a relative system (*quarter past three*). Testing five-minute intervals, Bock et al. showed that native speakers of Dutch prefer a relative system that uses the full and the half hour as reference points. 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 et al., the reference point (half hour or hour) changes at twenty and forty minutes past the full hour (*h:20* and *h:40*), respectively. In addition, the reference hour that is named in the utterance (*h* or *h+1*) changes at twenty minutes past the full hour. Figure 1 shows how the different utterance formats relate to the clock times.

Meeuwissen et al. (2003) measured Dutch clock time naming latencies for five-minute intervals on the digital clock (hour:minutes, *h:mm*), using the relative system. They found effects of the type of reference point (*h, half h+1, h+1*) and

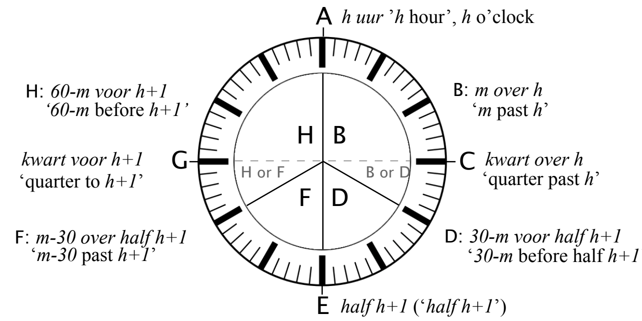

Figure 1: Preferred utterance formats for relative clock times in Dutch.

the distance (in minutes) between the minutes and their reference point. Utterances referring to the full hour (Figure 1, Format B) were faster than utterances referring to the next full hour (H), which in turn were faster than utterances referring to the half hour (D,F). The longer the distance (0, 5, or 10 minutes) between the minutes on the display and their reference point, the longer it took to initiate the utterance. In addition, Meeuwissen et al. found an interaction between distance and reference point, indicating that the distance effect is stronger for clock time naming latencies around the half hour and the next full hour.

The authors explain their findings by means of a procedural semantics for clock time naming, in which there are separate semantic procedures for determining the full hour (*h*) and the next full hour (*h+1*). Since clock times that refer to the next full hour always require the determination of both *h* and *h+1*, speech onset latencies are predicted to be longer for these times.

Furthermore, Meeuwissen et al. explain the interaction between reference point and distance as an effect of the numerical transformation that is required for minutes relative to the half hour and the next full hour, but not relative to the full hour (e.g., *h:20* becomes *ten before half h+1*, whereas *h:10* remains *ten past h*).

To summarize, Meeuwissen et al.'s model explains the speech onset latencies for clock time naming by means of three factors: (1) numerical transformation of the hour, (2) numerical transformation of the minutes, and (3) distance between the minutes to be transformed and their reference point.

While (1) and (2) seem rather straightforward explanations for an increase in reaction times, the third factor remains puzzling. Why should it take longer to say *ten before half two* than *five before half two*? Meeuwissen et al. do not specify the cognitive processes responsible for this effect. In fact, it is unclear what the cognitive equivalent to "distance in minutes" should be, weakening the viability of the model. If clock time naming is to be applied as a paradigm for testing phenomena in, for example, speech production, the underlying cognitive principles need to be specified in more detail. We will discuss three possible explanations that each specify a different aspect of transformation and distance.

To begin, one might think of the *physical* distance between the hands of an imaginary analog clock, where the small hand points to the hour and the big hand points to the minute. If one were to turn the big hand to its reference point (either the full hour or the half hour), it would take longer, the larger the distance in degrees. Most importantly, one would predict the effect of distance to be a linear function, for example $rt = degrees * x\ ms$. Such a linear effect is included implicitly in Meeuwissen et al.'s data plots, where the reaction times for the five-minute intervals that were tested in their experiment are connected by straight lines. The linear distance hypothesis implies that speech onset latencies for *all* minutes within an utterance format should show a linear increase in reaction time with distance.

Another possible explanation of the distance effect, that is related to the problem of minute transformation, involves the influence of mental arithmetics. Naming the clock times from a digital clock requires speakers to transform the number on the display. In order to generate the correct expression in the Dutch relative system, arithmetic problem solving is required. The required operations depend on the reference point: $0 + m$ for utterance format B, *30-m*(D), *m-30* (F), and *60-m* (H). The arithmetic hypothesis implies that the mental arithmetic required for transforming digital clock times can explain the pattern of speech onset latencies.

Another source of influence on the speech latencies for clock time naming might be found in the frequencies of the time expressions. Bock et al. (2003, p.683) suggest that time expressions are "a kind of non-figurative idiom, construction, or formula" (e.g., Kuiper, 1996), but show that nevertheless their linguistic formulation proceeds compositionally and in an incremental fashion. Accordingly, the general format of an idiom for relative clock times in Dutch is *x [prep] [half] y*, with *x* as the minute term, *[prep]* as an obligatory variable that can either be *voor* or *over* (before or past), *[half]* as an optional element, and *y* as the hour term. We assume that this abstract structure is not computed online, but rather accessed as a whole, and that its open slots are filled dependent on the actual time information. On the other hand, one might assume that frequently used time expressions are represented separately, for example by means of their own lexical entry whose minute slot has already been filled (e.g, "ten past y").

This suggests that the frequency of a time expression is related to its speech onset latency. Thus, the frequency hypothesis implies that the frequency distribution of time expressions can explain part of the variance in the speech onset latencies, and thus serves as a relevant factor in a model of clock time naming.

The goal of our present study is to assess the validity of these hypotheses for clock time naming. By identifying different factors that influence reaction times we will develop a more fine-grained model of clock time naming.

Experiment 1

The first experiment tested the hypothesis that speech onset latencies for digital clock times are a linear function of the physical distance between the minute hand and its reference point on an imaginary analog clock. Therefore we extended Meeuwissen et al.'s (2003) digital clock time naming paradigm: not only the standard five minute intervals were tested, but *all* minutes of the hour. The distance hypothesis predicts a linear increase of speech onset latencies with distance to reference point.

Method

Participants All participants in the present study were students of the University of Nijmegen and native speakers of Dutch who were paid for their participation. Twelve participants were tested in Experiment 1.

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 five-minute intervals (*standard times* in the following) 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'. They were explicitly instructed to use the nearest full or half hour as reference point for the formats B, D, F and H (the maximum distance being 14 minutes). Response latencies were measured by voice key.

At the beginning of each trial a fixation cross was presented for 500ms, followed by 150ms blank screen. Then the clock time stimulus was presented for 1000ms. Speech onset latencies were measured from clock presentation onset, with a deadline of 2500ms. A new trial was initiated 1500ms 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.

Analyses and Results All utterances were checked for erroneous or missing responses and disfluencies. Data from two participants were removed from the data set, because of more than ten percent errors. The results of Experiment 1 are shown in Figure 2. The dots in the top panel show the speech onset latencies per minute. The dashed lines connecting the minutes from 0 to 15 and from 45 to 59 show the latencies for filler times, which were included in order to increase the external validity of the task. The dashed line connecting all standard times accentuates the subset of five minute intervals tested by Meeuwissen et al. (2003). The plotted latencies are based on 94.75% of all data (6.25% being rejected because of errors, disfluencies, or latencies shorter than 350ms).

Given the continuous character of the *minute* variable, we analyzed these data using multilevel multiple regression models (Pinheiro & Bates, 2000), with the logarithm of the speech

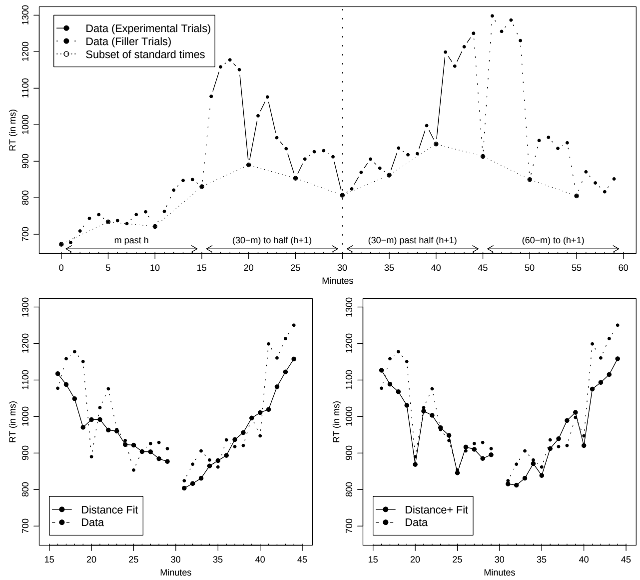

Figure 2: Speech onset latencies and fits of two different regression models for Experiment 1.

onset latencies as the dependent variable and participant as the error stratum.

Following Meeuwissen et al. (2003), we entered three types of predictor variables into the analyses: (1) magnitude information (the minutes), (2) length information (number of morphemes and number of phonemes of the minute term in the utterance), and (3) frequency (logarithm of the morpheme frequency for the minute term). In contrast to Meeuwissen et al., we preferred minute-term related variables over whole-term variables, as Dutch relative clock time expressions start with expressing the minute term. Moreover, combining the values for minutes and hours into one factor excludes the possibility of assessing separate and possibly independent contributions of hour and minute terms. To account for possible hour related effects, we included hour as a categorical variable in the analyses. Note that this standard set of variables does not include whole form frequency, as testing the contribution of frequency was a specific focus of Experiment 3. The standard set of predictor variables was used in the analyses of both Experiment 1 and 2, in addition to the sets of experiment-specific variables. In all analyses, we first entered the total set of variables as fixed effects and trial number as random effect, and constructed the best-fitting model based on stepwise model selection by exact AIC (Pinheiro & Bates, 2000).

As a test of replication, we first analyzed the subset of standard times. In addition to the variables mentioned above, we included a categorical variable representing the reference point (i.e., hour, half hour or next hour) and a categorical variable representing the distance to reference point (i.e., 0/15 for the cardinal times and either 5 or 10 for the remaining time points). The best-fitting model (all df=899) contains three predictors: the logarithm of the minute morpheme frequency (β=.179, t=7.923, p$<$.001), the reference hour (for the full hour β=-.174, t=-10.632, p$<$.001, for the next full hour: β=-.064, t=-3.478, p$<$.001), and the distance to reference point (for 5 minutes β=-.465, t=-7.739, p$<$.001, for 10 minutes: β=-.469, t=-7.795, p$<$.001). Note that in contrast to the estimates in Meeuwissen et al's work, the β estimated for the morpheme frequency effect is negative, indicating faster responses for utterances starting with high frequent morphemes.

To test the distance hypothesis, we selected the data ranging from 16 to 29 and 31 to 44 minutes. Two models were fitted to these data. The simpler model contained, apart from the standard set of variables, a continuous variable expressing the absolute distance in minutes to the half hour and a categorical variable indicating whether the time was past (31-44) or to (16-29) the half hour. This variable was included to allow separate effects for clockwise and anti-clockwise distance, and was therefore entered both as a main effect and as an interaction with distance. If the distance hypothesis is correct, this model should explain a reasonable amount of variance. However, as the top panel of Figure 2 shows, the response latencies for the standard times seems to be shorter than those for the non-standard times. Therefore, we tested an additional, slightly more complex model to the data. This model contains an additional binary variable indicating whether the item is one of the standard times. Whether the addition of this parameter resulted in a significant improvement of fit of the model, compared to the loss of degrees of freedom, was tested by means of a likelihood ratio statistic.

The best fitting simple model (all df=2022) contains a non-significant parameter for the logarithm of the morpheme frequency, a significant parameter for the number of phonemes (β=.025, t=3.109, p=.002) and a set of non-significant parameters for the factorial hour variable. Apart from these parameters, the model also contains a parameter expressing an effect of distance (β = .027, t=8.980, p$<$.001, indicating that each increase in distance is associated with an increase of .027 in log(RT)), a variable expressing a main effect of *before* versus *past* the half hour (β=.101, t=4.349, p$<$.001, indicating increased latencies before the half hour), and a variable expressing the interaction between distance and before/past (β=-.011, t=-3.999, p$<$.001, indicating a reduced increase in latency for distance past the half hour).

Starting from this simpler model, we added a binary variable representing standard times, which we allowed to interact with before/past and distance. The best-fitting complex model is very similar to the simpler model; the main difference is a significant parameter for the interaction between distance and standard times (β=-.016, t(2020)=-2.686, p=.007). This indicates that the standard times are faster than the interjacent time points, and that latencies are shorter for the 10 than for the 5 minute distance. This second model fitted the data significantly better than the first model (LR=64.68, p$<$.0001).

The fit of both models is depicted in the lower two panels of Figure 2. The bottom-left panel shows the fit of the simple model, completely missing the effect of standard times. The bottom-right panel shows the significantly better fit of the more complex model. The analyses show that, although distance explains subsets of the data reasonably well, a model relying solely on distance can easily be augmented with additional variables. This suggests that the standard clock times are processed differently from non-standard clock times. This might be caused by differences in arithmetic processing (which will be tested in Experiment 2), or by differences in the underlying representation of the expression.Bock et al., 2003).

Experiment 2

The second Experiment was designed to test the hypothesis that the difficulty of the mental arithmetics required during clock time naming influences the naming latencies. We measured speech onset latencies for arithmetic problems that resemble the assumed operations during clock time naming.

In order to allow a comparison with the reaction times found for clock time naming in Experiment 1, problems were constructed such that all time points of the hour were reflected. In addition, the time intervals from *h:16* to *h:19* and from *h:41* to *h:44* were represented by two different formats (B/D, and H/F, respectively), reflecting the fact that Dutch offers two alternative reference points for these time points. Consequently, two different arithmetic operations could potentially explain the reaction time data for these variable sets.

Method

64 different arithmetic problems were constructed, in four different conditions. The design reflected the distribution of alternative formats for clock times in Dutch, as depicted in Figure 1, excluding the cardinal times. 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 answer with an initial *is*, as in *(thirty minus seven) is twenty-three*. There were 18 items each for the Formats B and H, and 14 items each for the Formats D and F. Each item was repeated four times, resulting in a total of 256 trials. 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 Experiment 1. Fifteen participants were tested.

Analyses and Results The speech onset latencies of Experiment 2 are shown in Figure 3. We fitted a model to these data

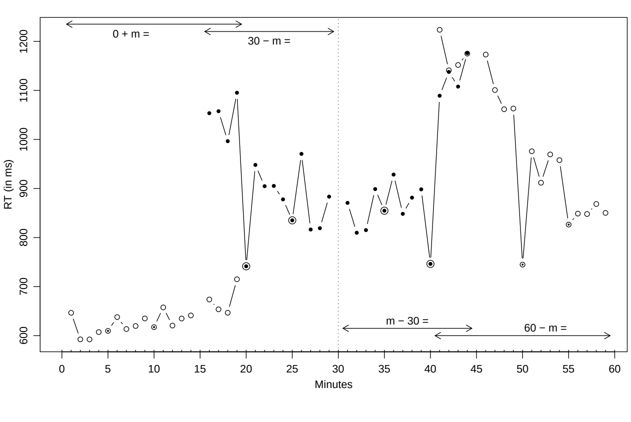

Figure 3: Speech onset latencies of Experiment 2.

using similar procedures as described for Experiment 1. The variables entered in the regression are (1) a numerical representation of the solution (similar to distance in Experiment 1), (2) a categorical representation of the reference point, and (3) a binary variable that distinguishes the eight standard times (depicted with an additional grey circle in Figure 3) from the remaining time points. All interactions of these factors were included. This initial model also proved to be the best fitting model. However, for the sake of simplicity, and because the estimated parameters are highly similar to the full model, we will discuss a more simple model in which the type of arithmetic problem is represented as a two level factor, either simple ("0+") or complex ("30-", "-30", or "60-"). The significant estimated parameters (all df=3168) are the magnitude of the solution (β=.006, t=4.426, p$<$.001), a parameter for problem type (simple vs. complex, β=.282, t=14.595, p$<$.001), as well as interactions between magnitude and problem type (β=.015, t=8.331, p$<$.001) and between standard time, magnitude and problem type (β=-.037, t(3168)=-3.423, p$<$.001).

These results show that both a larger difference between the two operants in the arithmetic problem and the complexity of the problem lead to an increase in latencies. However, while in the more complex conditions the default effect is an increase of latencies with solution magnitude, the effect for the problems that involve five and ten is *reversed*. That is, solving 40-30 or 30-20 is significantly faster than solving 35-30 or 30-25. Thus, the effect of mental arithmetics is diametrically opposed to the effect for standard times as reported in Experiment 1 and Meeuwissen et al's (2003) study. In other words, mental arithmetics can explain the distance effect as an effect of magnitude of solution. However, this explanation does not hold for the subset of time points that were originally tested by Meeuwissen et al. (2003).

Frequency estimates

Experiment 2 has shown that the time required for calculations does not explain the full pattern of results in Experiment 1. As was discussed previously, the frequency of the clock time expressions might explain the difference between standard times and non-standard clock times.

Following our reasoning in the introduction, we consider clock time expressions as compositional idioms or formulas. As frequency estimates for idioms are not simply a function of the frequency of their component parts (Sprenger, 2003), we assessed the frequencies of the initial clusters of Dutch relative clock time expressions (averaging across hours).

Method

We determined the frequencies of 103 clock time clusters of the format *[x (alphabetic) [prep] [half]]* by subjecting them to a *Google* search of Dutch webpages. The number of unique webpages 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. The range of queries covered all clock times in the Formats A-D, as shown in Figure 1). However, the range of the Formats B and H was extended: Format B was used for queries from 1 to 25 minutes past the hour, Format H for queries from 25 to 1 minute before the next hour. This procedure allowed to compare the frequencies of possible alternative utterance formats for the variable sets.

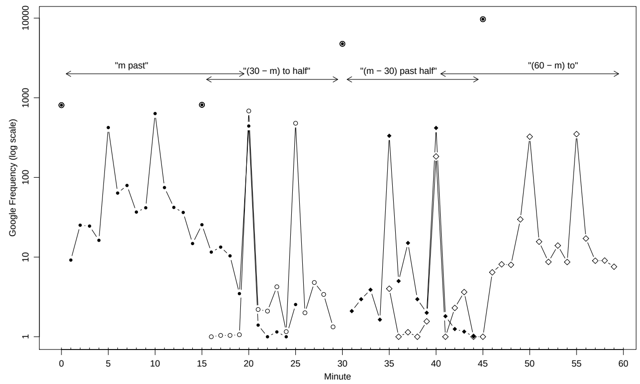

Figure 4: Frequencies of Dutch clock time expressions.

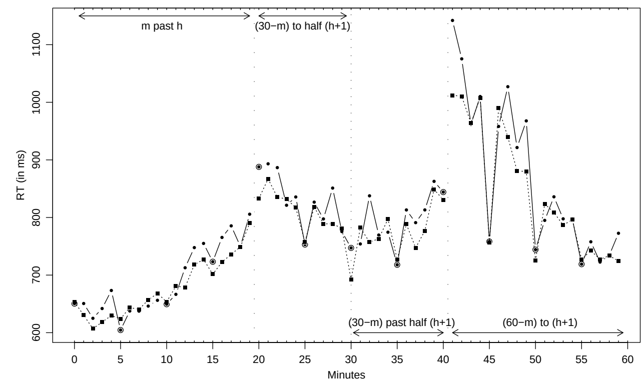

Figure 5: Speech onset latencies (continuous line) and model predictions (dashed line) for Experiment 3

Results and Discussion

The results of the frequency counts are shown in Figure 4. The cardinal times (represented by circles with an additional ring) are more frequent than the other times. Within the set of non-cardinal times, the standard times are in turn more frequent than the non-standard times. For each of the formats it can be observed that the frequency first increases with distance to reference point, and then decreases again.

A more specific effect can be observed for the minutes that have two alternative reference points (16-25, 35-44). Within these ranges, the frequency favors the B and H formats for time points closer to the hour, and format D and F for time points closer to the half hour. However, this preference is not absolute. For the 20 and 40 minute time points, the frequencies of the *relative to the half hour* formats (D, F) are only slightly higher than the frequencies of the *relative to the (next) full hour* formats (B, H). This is most evident in the relative proportions of the preferred formats. The only time points for which this proportion is below .75 are 20 (P(D)=.52), 21 (P(D)=.70), 39 (P(F)=.61) and 40 (P(F)=.54, all proportions are calculated from the logarithm of the frequency). Note that the frequencies of five minute distances are lower than those of ten minute distances (as operationalized in Experiment 1). Assuming a negative correlation between frequency and speech onset latencies, this frequency difference is the opposite of what would be necessary to explain the distance effect. However, given that there are two alternative high frequent formats for the ten minute distance, the increase in speech onset latencies that was observed both by Meeuwissen et al. and in Experiment 1 might be explained in terms of competition between these two formats. This assumption will be tested in Experiment 3.

Experiment 3

In the third experiment we measured speech onset latencies for clock times, using the preferred clock time format for each time point (as determined by the frequency estimates).

Method

Materials, Design and Procedure were identical to Experiment 1, except for the instruction. Participants were asked to use clock time format B (instead of D) if the minutes ranged from 16 to 19, and format H (instead of F) if they ranged from 26 to 44 (variable sets). Thirteen participants were tested.

Analyses and Results The speech onset latencies of Experiment 3 are shown in Figure 5. Figure 4 shows the data of Experiment 3 in continuous lines with circle markers, and the regression model fit in dashed lines with square markers. Compared with Experiment 1, the main differences are the speedup in overall response latency and the faster speech onset latencies for the 16-19 and 41-44 minute times, both indicating that this experiment tested a more natural setup.

To test the hypotheses discussed in the introduction of this paper, we entered sets of factors representing distance, arithmetic difficulty and cluster frequencies in the regression equation and assessed the optimal model. For distance, we entered minutes in the equation both as a main effect and as interaction with utterance format (in which we collapsed the formats for 0, 15, 30 and 45 minutes). By means of the interaction, distance effects per utterance format can be assessed. For arithmetic difficulty, we entered four factors into the equation, each representing a different type of arithmetic problem. This is similar to the interaction used in the analyses of Experiment 2, but as some arithmetic problems have not be tested in Experiment 2 (e.g., *30-30*) we had to divide this predictor into four separate factors. For the cluster frequencies, we included (1) the logarithm of the cluster frequency in interaction with a binary factor representing cardinal times, (2) a factor representing the probability of preferred representation and (3) the interaction between the logarithm of the frequency and the probability for the minutes in the 16:25 and 35:44 range. Apart from these factors, we also included the standard factors as discussed earlier. We again assessed the best fitting model using stepwise model selection.

The resulting model (see Figure 5) contains terms from all three hypotheses. The distance hypothesis is represented by a significant interaction between minute and utterance format (β_B=.006, β_D=-.015, β_F=-.003, β_H=-.015, $F(4,4504)=11.05$, $p<.0001$). The positive parameter for format B and the negative parameters for the other formats indicate a significant distance effect (which is not due to the standard times, as including the standard times in the equation does not significantly change the fit). The arithmetic hypothesis is represented by parameters for arithmetic latencies in the four main utterance formats (β_{B}=2.289, t=-1.733, p=.083, β_{D}=-2.098, t=-1.590, p=.112, β_{F}=3.174, t=6.928, p<.001, β_{H}=3.223, t=7.080, p<.001.[1] The frequency hypothesis[2] is represented in the final model by parameters for probability of utterance format, the logarithm of the frequency, as well as interactions between frequency and cardinal times and between frequency and probability. The parameters are estimated at -.172 (t=-3.944, p<.001) for probability of utterance, at -.013 (t=-3.492, p<.001) for the logarithm of the utterance frequency, at .016 (t=3.301, p<.001) for interaction between logarithm of the utterance and cardinal times, and at -.009 (t=-2.022, p=.043) for the interaction between probability and frequency for the range of minutes between 16 and 25 and 35 and 44. The most interesting of these parameters is the probability, indicating that clock times for which multiple formats compete have longer speech onset latencies. The best fitting model contained, apart from these hypothesis-related parameters, only non-significant parameters.

Discussion

Clock time naming is a new paradigm in the speech production literature that employs a relatively simple task to investigate the production of complex utterances. However, the complexity of the utterances and the nature of the task demand a precise model of the cognitive factors involved.

In the present study we sought to test and further specify Meeuwissen et al.'s (2003) model of clock time naming. As discussed in the introduction, we tried to determine the cognitive factors underlying the distance effect. Therefore we extended the paradigm to all minutes of the hour. Three hypotheses were tested: (1) the distance hypothesis, which predicted an increase of speech onset latencies with increasing distance to reference point, (2) the arithmetic hypothesis, which predicted that the pattern of speech onset latencies can be explained by the difficulty of the underlying arithmetic operations, and (3) the frequency hypothesis, which predicted a significant influence of the frequency of clock time expression on speech latencies.

The results of Experiment 1 show that the extended clock time naming paradigm reveals a more complex pattern of speech onset latencies, thereby supporting our hypothesis that the Meeuwissen et al. model is underspecified. Most importantly, a clear difference between speech onset latencies for the standard times and the remaining time points appears, with latencies for standard times being faster. A simple regression model that includes distance to reference point as a factor does not result in a proper fit; a model including the standard times is siginificantly better. However, this second model still does not accurately predict the standard times.

The results of Experiment 2 showed that the time required for mental arithmetics depends on the type of arithmetic problem, as well as on the magnitude of its result. The latter result resembles the distance factor in clock time naming. However, closer inspection of the data revealed that the magnitude effect does *not* hold for the arithmetic problems that correspond to the standard times. In other words, while mental arithmetics is a significant factor in clock time naming that can serve as a probable explanation for the distance effect, it also demonstrates that a complete model has to contain additional factors.

Following Bock et al.'s assumption that clock time names are idioms, and assuming with Sprenger (2003) that idiom frequency is not a simple function of its component parts, we conducted a corpus study on clock time expression frequencies. The results indicated that some time expressions are used much more often than others, which is intuitive for cardinal times, but also holds for the standard times. Thus, frequency of usage can partly explain the difference between standard times and non-standard times. In addition, the frequency counts revealed that for some time points there are preferred and dispreferred formats. While for most time points the preference is relatively unambiguous, the time points *h:20* and *h:40* show high frequencies for both utterance formats. The resulting competition of utterance formats can explain the longer reaction times for these time points that were observed by Meeuwissen et al. and in Experiment 1.

In our third experiment we measured clock time naming latencies, using only the preferred clock time formats. The regression model for these data showing the best fit includes arithmetics, frequency and distance as a factor. However, Experiment 1 showed that distance without frequency and arithmetics does not yield a good fit. Moreover, the significant contribution of preference in this model supports the idea of separate idiom representations for high frequent time points.

To sum up, the results of this study suggest that a complete model of clock time naming needs to take into account the type of the mental operations that underly numerical transformation, as well as the frequency and possible idiomaticity of the expressions involved.

[1] Although not all parameters are significant at the .05 level, excluding these parameters decreases the fit (LR=5.87, p=.053). Moreover, as splitting the arithemtic latency into four factors was because of technical reasons, removing just two of these parameters would be illogical.

[2] Note that Meeuwissen et al. also included whole-form frequency in some regression analyses. However, apart from mentioning the estimated parameters, they do not elaborate on the role of frequency in clock time naming.

References

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

Kuiper, K. (1996). *Smooth talkers: the linguistic performance of auctioneers and sportscasters*. Mahwah, NJ: Erlbaum.

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

Pinheiro, J. C. & Bates, D. M. (2000). *Mixed-Effects Models in S and S-Plus*. New York: Springer.

Sprenger, S. A. (2003). *Fixed expressions and the production of idioms*. Doctoral dissertation, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.