A TIME-TO-EVENT ANALYSIS OF AUDITORY LEXICAL DECISION LATENCIES

Peter Hendrix, Ching Chu Sun

Eberhard Karls Universität Tübingen
peter.hendrix@gmail.com

ABSTRACT

We report the results of an analysis of the auditory lexical decision latencies in the Massive Auditory Lexical Decision database (MALD) [9] using a statistical technique for time-to-event analysis: the piece-wise exponential additive mixed models (PAMM) [2, 1, 3]. The PAMM models the probability of an instantaneous response at each point in time, rather than the response times themselves. The analysis revealed an increased instantaneous probability of a response for high frequency words, as well as for words from sparse phonological neighborhoods. These effects were most prominent during the early stages of the decision making process, but remained significant throughout large parts of the response window. Furthermore, we observed a more transient early effect of the temporal distance to the uniqueness point. The PAMM analysis of the MALD data thus provides more insight into the temporal dynamics of lexical processing in the auditory lexical decision task.

Keywords: auditory lexical decision, time-to-event analysis, piece-wise exponential additive mixed model, uniqueness point, phonological neighborhood density

1. INTRODUCTION

Analyses of experimental data in phonetic psycholinguistics typically focus on the mean of the response variable distribution. A least squares regression of the response times in an auditory lexical decision experiment, for instance, estimates the conditional mean of the response time distribution given one or more lexical predictors, such as the length or frequency of a word. Effects of lexical predictors, however, need not be constant over the response time distribution. The effects of some predictors may primarily influence short reaction times, whereas the effects of other predictors may be more prominent for long reaction times. Furthermore, due to the temporal nature of the speech signal, the values of lexical or acoustic predictors themselves may vary as a function of time.

Further insight into the temporal dynamics of predictor effects in behavioral experiments helps further our understanding of the language processing system. Distributional analyses help provide such insight. Here, we present a distributional analysis of auditory lexical decision data using a statistical model that is based on the principles of time-to-event analysis: the piece-wise exponential additive mixed model (PAMM) [2, 1, 3]. The PAMM allows for an investigation of non-linear effects of both time-constant and time-varying predictors as they develop over time. Here, we use a PAMM to gain further insight into the temporal development of the effects of word frequency, phonological neighborhood density, and the phonological uniqueness point on the behavioral responses in the Massive Auditory Lexical Decision (MALD) database [9].

2. METHODS

From the Massive Auditory Lexical Decision database (MALD [9]) we extracted average lexical decision latencies for all words with at least one correct response. This resulted in a set of 26,520 words. We analyze the average lexical decision latencies for these words using a PAMM. Rather than the response times themselves, the PAMM models the instantaneous probability of a response throughout the response window. One advantage of this approach is that predictors need not be constant over the response time distribution. Here, we investigate the effects of two time-varying predictors: *time since offset* and *time since UP*. Furthermore, we entered two time-constant predictors into the PAMM analysis: *frequency* and the phonological neighborhood density measure *PLD*. We describe these time-constant and time-varying predictors in more detail below.

The MALD database provides the sound file and phoneme level segmentation for each pronunciation. On the basis of these data, we calculated the acoustic duration and the acoustic uniqueness point (henceforth UP) for each of the words under investigation. The acoustic duration is the time from stimulus onset to the temporal offset of the final phoneme of a word in the acoustic signal. Similarly, we defined the acoustic UP as the temporal endpoint of the phoneme that distinguishes a word from all other words, including the members of a word's morphological family.

As noted above, PAMMs offer the opportunity to include time-varying predictors into the analysis. Consequently, we did not enter the acoustic duration and acoustic UP into the analyses directly. Instead, we defined the corresponding predictors *time since offset* and *time since UP* as the temporal distance between the current point in time on the one hand and the the offset of the acoustic signal and the UP on the other hand. At time $t = 400$ ms, for instance, the values of *time since offset* and *time since UP* for a word with an acoustic duration of 500 ms and a UP of 350 ms are -100 ms and 50 ms, respectively.

We furthermore investigated the effects of two time-constant predictors: *frequency* and *PLD*. We defined *frequency* as the frequency of the orthographic word form in the SUBTLEX-US corpus [4]. As shown by Tucker et al. [9], the explanatory power of the orthographic frequencies from SUBTLEX-US for the auditory lexical decision data in the MALD database is highly competitive as compared to phonological frequencies from corpora of spoken speech, such as the spoken language subset of the Corpus of Contemporary American English (COCA [5]). The SUBTLEX-US frequency counts were log-transformed prior to analysis to remove a rightward skew from the frequency distribution.

PLD is a measure of phonological neighborhood density that is based on the phonological Levenshtein distance between words. The phonological Levenshtein distance between two words is the total number of deletions, additions, or substitutions that are necessary to convert the phonological form of one word into the phonological form of another word [7]. Technically, the *PLD* measure used here is defined as the average phone-level Levenshtein distance between a word and all other words in an adapted version of the CMU Pronouncing Dictionary [10, 9]. Prior to analysis, we applied an inverse transformation ($f(x) = \frac{-1}{x}$) to the phonological Levenshtein distances to increase the symmetry of the *PLD* distribution.

3. ANALYSIS

As noted above, we analyzed the auditory lexical decision latencies in the MALD database with a statistical technique from time-to-event analysis: the PAMM. [2, 1, 3]. The PAMM is framed within the context of the generalized additive mixed-effect model (GAMM) [11, 12]) and makes it possible to uncover non-linear predictor effects that vary in a non-linear fashion as a function of time. Furthermore, the values of predictors themselves need not be constant over time. Recently, Hendrix [6] adopted PAMMs to investigate non-linear time-varying predictor effects in the visual lexical decision task.

Rather than the response time itself, the response variable in a PAMM is the (log of the) instantaneous hazard rate: the probability that an event of interest occurs at time t, provided that it did not occur prior to time t. The event of interest in the current study is the "word or non-word" decision in the auditory lexical decision task. Conceptually, the PAMM is an extension of the piece-wise exponential model (PEM) in the sense that the (log of the) instantaneous hazard rate is estimated in a piece-wise fashion for each of a number of time intervals in the response window (i.e., the time window in which lexical decisions come in). Technically, for all time points in the interval $j := (\kappa_{j-1}, \kappa_j]$, the (log of the) hazard function $\lambda(t|\mathbf{x}_i)$ given the predictor values $\mathbf{x}_i$ for stimulus i is defined as:

$$(1) \quad \log(\lambda(t|\mathbf{x}_i)) = \log \lambda_0(t_j) + \sum_{k=1}^{p} f_k(x_{i,k}, t_j)$$

where $\lambda_0(t_j)$ is the baseline hazard for time interval j, and $f_k(x_{i,k}, t_j)$ are smooth functions for predictors $k \in 1, \ldots, p \ \forall \ t \in j$. Note that random effect structures may be specified in a PAMM as well. No random effects were included in the current analysis, however. We therefore omitted the specification of random effect structures from Equation 1.

The baseline hazard is the (log of the) overall probability of a response as it evolves over time and is modelled through the model intercept and a smooth over time (i.e., through a `s(time)` term). Predictor effects are adjustments to this baseline hazard. Here, we estimate time-constant effects of predictors through predictor smooths (i.e., `s(predictor)` terms), whereas we allowed for time-varying predictor effects by including tensor product interactions between time and predictor (i.e., `ti(time,predictor)` terms; see Wood (2017) [13] for more details). Predictor smooths as well as time by predictor interactions were limited to fourth order non-linearities to ensure interpretability of the results. Although it is possible to model three-way interactions between time and two predictors in a PAMM, we refrained from including such interactions in the analysis for easy of interpretation. For each predictor, predictor outliers further than 3 standard deviations from the predictor mean were removed prior to analysis.

4. RESULTS

The results of the PAMM analysis of the auditory lexical decision latencies in the MALD database are presented in Table 1. As can be seen in Table 1, both the model intercept (β = -4.678, $p < 0.001$) and the smooth of *time* (χ^2 = 1852.250, $p < 0.001$) were significantly different from zero. The resulting (log) baseline hazard is presented in the left panel of Figure 1. The probability of an instantaneous response is initially low and increases as a function of time, most prominently so between the start of the analysis window (640 ms after stimulus onset) and 800 ms after stimulus onset. As noted by Hendrix [6], this functional shape of the baseline hazard function is typical for response time distributions.

We observed a significant main effect of *time since offset* (χ^2 = 17.240, $p < 0.001$) as well. The effect of *time since offset*, however, strongly interacts with *time* (χ^2 = 1016.438, $p < 0.001$). The effect of *time since offset* is visualized in the right panel of Figure 1. Time is on the x-axis, whereas *time since offset* is on the y-axis. The z-axis represents the adjustment to the (log) baseline hazard as a function of *time* and *time since offset*, with warmer colors representing a higher probability of an instantaneous response.

The contour plot for *time since offset* includes the partial main effect of *time since offset* and the partial *time* by *time since offset* interaction, but excludes the partial main effect of *time*. The partial main effect of *time* is excluded for ease of interpretation and is omitted in all subsequent *time* by predictor contour plots as well. As expected, the probability of an instantaneous response is lower when the offset of the acoustic signal has not yet been reached (i.e., for negative values of *time since offset*).

We furthermore observed a significant main effect

Table 1: Results for a PAMM fit to the auditory lexical decision latencies in the MALD database [9]. Provided are β coefficients and p-values for parametric terms, as well as χ^2-values and p-values for smooth terms.

parametric terms	β	p**-value**
intercept	-4.678	<0.001
smooth terms	χ^2**-value**	p**-value**
time	1852.250	<0.001
time since offset	17.240	<0.001
time by *time since offset*	1016.438	<0.001
frequency	445.388	<0.001
time by *frequency*	94.399	<0.001
PLD	78.292	<0.001
time by *PLD*	31.985	<0.001
time since UP	12.963	<0.001
time by *time since UP*	71.322	<0.001

of *frequency* (χ^2 = 445.388, $p < 0.001$) and a significant *time* by *frequency* interaction (χ^2 = 94.399, $p < 0.001$). The effect of *frequency* is presented in the left panel of Figure 2. The probability of an instantaneous response is higher for high frequency words as compared to low frequency words. The effect of *frequency* is most prominent during the early stages of the decision making process, but remains significant throughout a large part of the response window. Indeed, the effect of *frequency* last reaches significance at 1,344 ms after stimulus onset, at which point in time no less than 96.44% of the words have been responded to. Word frequency thus continues to influence the decision making process during later stages of the response window.

In addition to the effect of *frequency*, the PAMM analysis revealed significant main effect (χ^2 = 78.292, $p < 0.001$), as well as a significant interaction with *time* (χ^2 = 31.985, $p < 0.001$) for *(inv) PLD*. The effect of *PLD* is presented in the middle panel of Figure 2. Consistent with the longer lexical decision latencies for words from dense phonological neighborhoods reported in earlier studies (see

Figure 1: Left panel: (log) baseline hazard. Right panel: effect of *time since offset*.

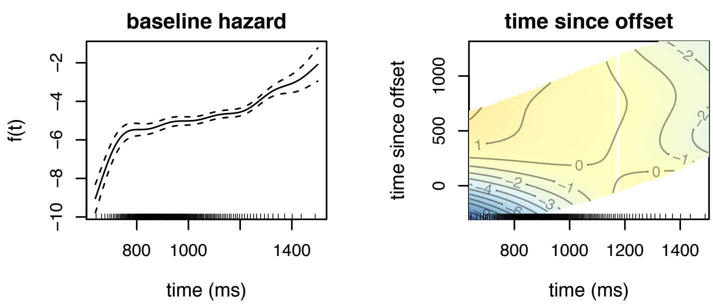
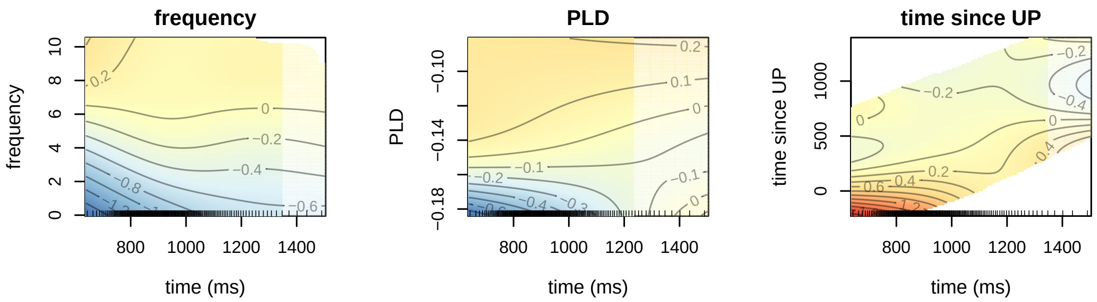


Figure 2: Effects of *frequency* (left panel), *PLD* (middle panel), and *time since offset* (right panel).



e.g., [8]), the probability of an instantaneous response is lower for words from dense phonological neighborhoods (i.e., for low values of *PLD*). As was the case for the effect of *frequency*, the effect of *PLD* is most prominent during the earlier stages of the decision making process, but remains significant until 92.84% of the words have been responded to (i.e., until 1,236 ms after stimulus onset).

Finally, both the main effect ($\chi^2 = 17.240$, $p < 0.001$) and the interaction with time ($\chi^2 = 1016.438$, $p < 0.001$) were significant for *time since UP*. The instantaneous probability of a response is higher when the acoustic uniqueness point is more recent (i.e., for smaller positive values of *time since UP*). Participants are thus more likely to respond when the acoustic signal of a word was recently distinguished from the acoustic signal for all other words. In comparison with the effects of *frequency* and *PLD*, the effect of *time since UP* is more transient in nature with large effect sizes at the start of the response window and much smaller effect sizes during later stages of the decision making process.

5. DISCUSSION

We reported the results of a time-to-event analysis of the auditory lexical decision latencies in the MALD database [9] using a PAMM [2, 1, 3]). We observed a time-varying effect of the temporal distance to the uniqueness point (i.e., the point in time at which a word can be distinguished from all other words). As expected, the probability of an instantaneous response was higher when the uniqueness point was more recent. This effect, however, was much more prominent during the earlier stages of the response window than at later points in time. During the earlier stages of the response window, responses come in for words that are easily identified as real words. The current results indicate that responses to such words tend to come in soon after or even before the point in time at which a word can be distinguished from all other words.

For words that cannot be identified as easily the temporal location of the uniqueness point is less relevant. Instead, listeners resort to more static sources of information, such as the frequency or phonological neighborhood density of a word. Consistent with previous findings, the probability of an instantaneous response was higher for high frequency words and for words from sparse phonological neighborhoods. Although the effects frequency and phonological neighborhood density were most prominent during the early stages of the response window, however, both predictors continued to show robust and qualitatively consistent effects throughout the response window. The information provided by these lexical-distributional measures thus remains relevant for the decision making process throughout the analysis window.

As a first exploration of PAMMs in the context of speech perception, the work reported here focuses on the effects of a limited number of lexical predictors on the probability of an instantaneous response in the auditory lexical decision task. For these predictors, however, the current analysis helped gain further insight the temporal development of the nonlinear effects of the predictors on lexical processing in the auditory domain. The (relative) timing of predictor effects is crucial for the development of psycholinguistic theories and models of speech perception. The results reported here thus suggest that PAMMs have the potential to uncover valuable information that is not available through more traditional analysis techniques.

6. REFERENCES

[1] Bender, A., Groll, A., Scheipl, F. 2018. A generalized additive model approach to time-to-event analysis. *Statistical Modelling* 18, 299–321.

[2] Bender, A., Scheipl, F. 2018. pammtools: Piecewise exponential additive mixed modeling tools. https://arxiv.org/pdf/1806.01042.pdf.

[3] Bender, A., Scheipl, F., Hartl, W., Day, A. G., Küchenhoff, H. 2018. Penalized estimation of complex, non-linear exposure-lag-response associations. *Biostatistics*. https://doi.org/10.1093/biostatistics/kxy003.

[4] Brysbaert, M., New, B., Keuleers, E. 2012. Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods* 44(4), 991–997.

[5] Davies, M. 2009. The 385+ million word Corpus of Contemporary American English (1990-2008+): design, architecture, and linguistic insights. *International Journal of Corpus Linguistics* 14(2), 159–190.

[6] Hendrix, P. 2018. A word or two about nonwords. *Manuscript under revision for Journal of Experimental Psychology: Learning, Memory, and Cognition*.

[7] Levenshtein, V. 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics – Doklady* 10, 707–710.

[8] Luce, P. A., Pisoni, D. B. 1998. Recognizing spoken words: The neighborhood activation model. *Ear & Hearing* 19, 1–36.

[9] Tucker, B. V., Brenner, D., Danielson, D. K., Kelley, M. C., Nenadić, F., Sims, M. 2018. The Massive Auditory Lexical Decision (MALD) database. *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-1056-1.

[10] Weide, R. 2015. The Carnegie Mellon Pronouncing Dictionary [cmudict. 0.6]. http://www.speech.cs.cmu.edu/cgi-bin/cmudict.

[11] Wood, S. N. 2006. *Generalized Additive Models*. New York: Chapman & Hall/CRC.

[12] Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)* 73(1), 3–36.

[13] Wood, S. N. 2017. *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC 2 edition.