# BLC article (stress, natives, late advanced learners and interpreters)

## Overview

This document contains updates to the statistical analysis for BLC article (3rd revisions). Last udpated on 2019-05-24. The results section can be copied and pasted into the corresponding google doc. The tables can also be copy and pasted where appropriate.

## Main changes

**Participants**

The updated analysis includes more participants in the interpreter (IN) group and a few different participants in the monolingual (M) and advanced learner non-interpreter (NIN) groups.

**Analyses**

The analysis has been simplified to in 2 major ways: by **removing the t-test and GLMM analyses**. The motivation behind using the t-tests was to see if participants fixated on targets above chance levels (50%) at the offset of the first syllable of the target items. The problem with this analysis is that is reduced a lot of data to approximately 50 data points (and 10ish in the IN group). This forced us to run multiple tests and correct for family-wise error using a Bonferroni adjustment (and killing our power). This was particularly troubling becuase it was apparent that there was a stress effect (oxytones consistently had higher fixation rates), but we refrained from testing syllable structure to avoid (1) futher correcting alpha and (2) harking. In place of the t-tests I have included model estimates from the growth curve analysis at the target syllable offset ± SE. This is not a formal test against chance, but is clearly more reliable than the t-tests. It also includes the 50ms bin, which I believe will make one of the reviewers happy. The downside is that it inevitably includes some acoustic information from slightly after the initial syllable. We don’t know how much, but we can examine the time course to see how close taret fixations are to chance before and after the relvant bin (bin #4, 4 x 50 = 200 ms). With regard to the GLMMs, I believe we can obtain the same information (and more) from the GCA.

## Main findings

The most important change regarding the findings is the **significant effect of lexical stress**. I believe the power increase due to additional participants is the principle explanation for this. I was able to get the most complex (maximal) random effects structure to converge. The story seems to be the following: native speakers anticipate target suffixes in all conditions, though certain conditions seem to facilitate processing. Specifically, if we consider paroxytone words with open syllables to be the default (the most common syllable in Spanish), we see earlier target fixations with the addition of the coda and with a shift of stress to the final syllable, but the effects are not compounding. For example, an unstressed penult is associated more looks to target, but the addition of a coda doesn’t help that much more.

# Plots

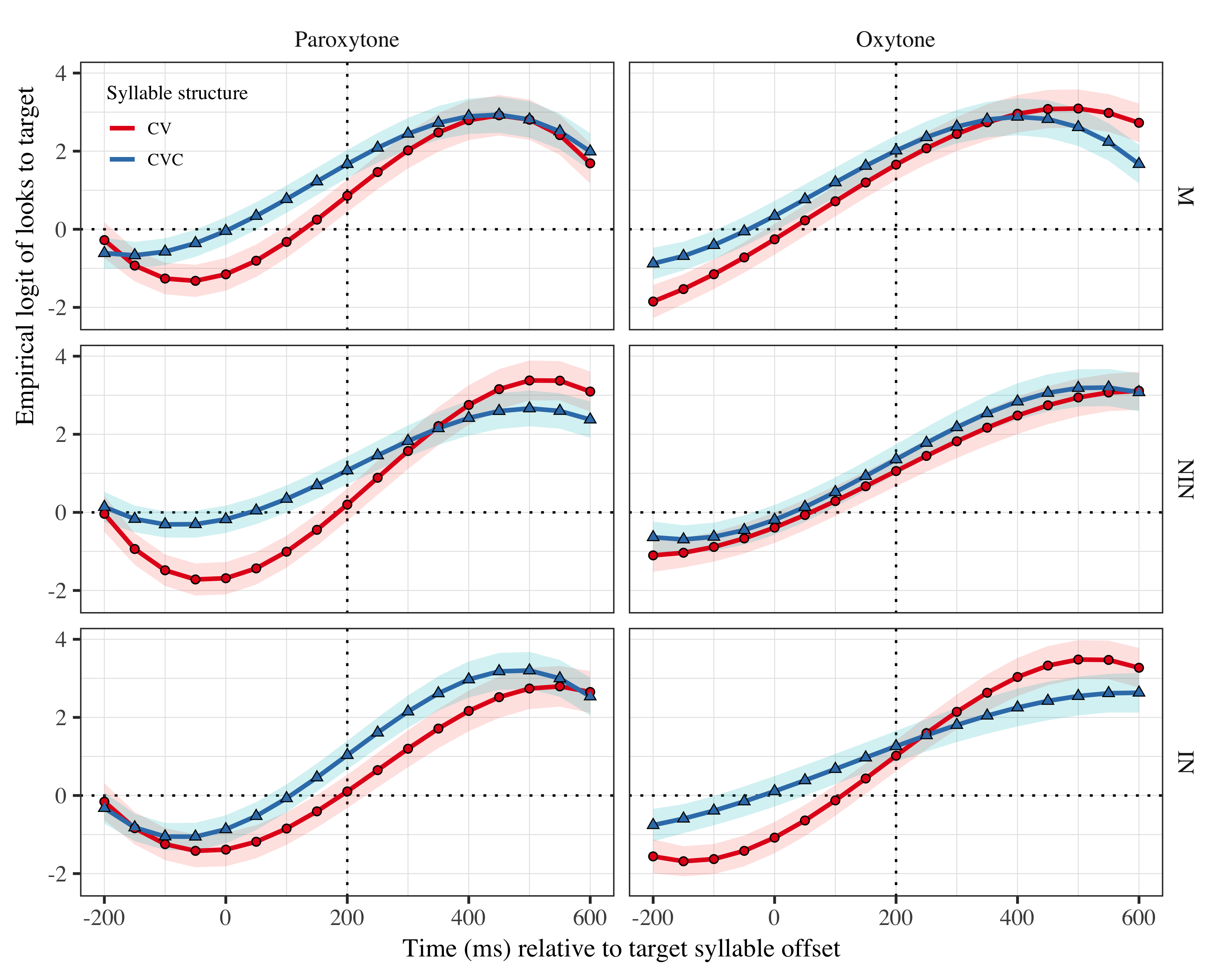


Figure 1: Growth curve estimates of target fixations as a function of lexical stress and syllable structure for each group during the analysis window. Symbols and lines represent model estimates, and the transparent ribbons represents ±SE. Empirical logit values on y-axis correspond to proportions of 0.12 0.50 0.88 0.98. The horizontal dotted line represents the 50% probability of fixating on the target. The vertical dotted line indicates 200 ms after the offset of the target syllable.

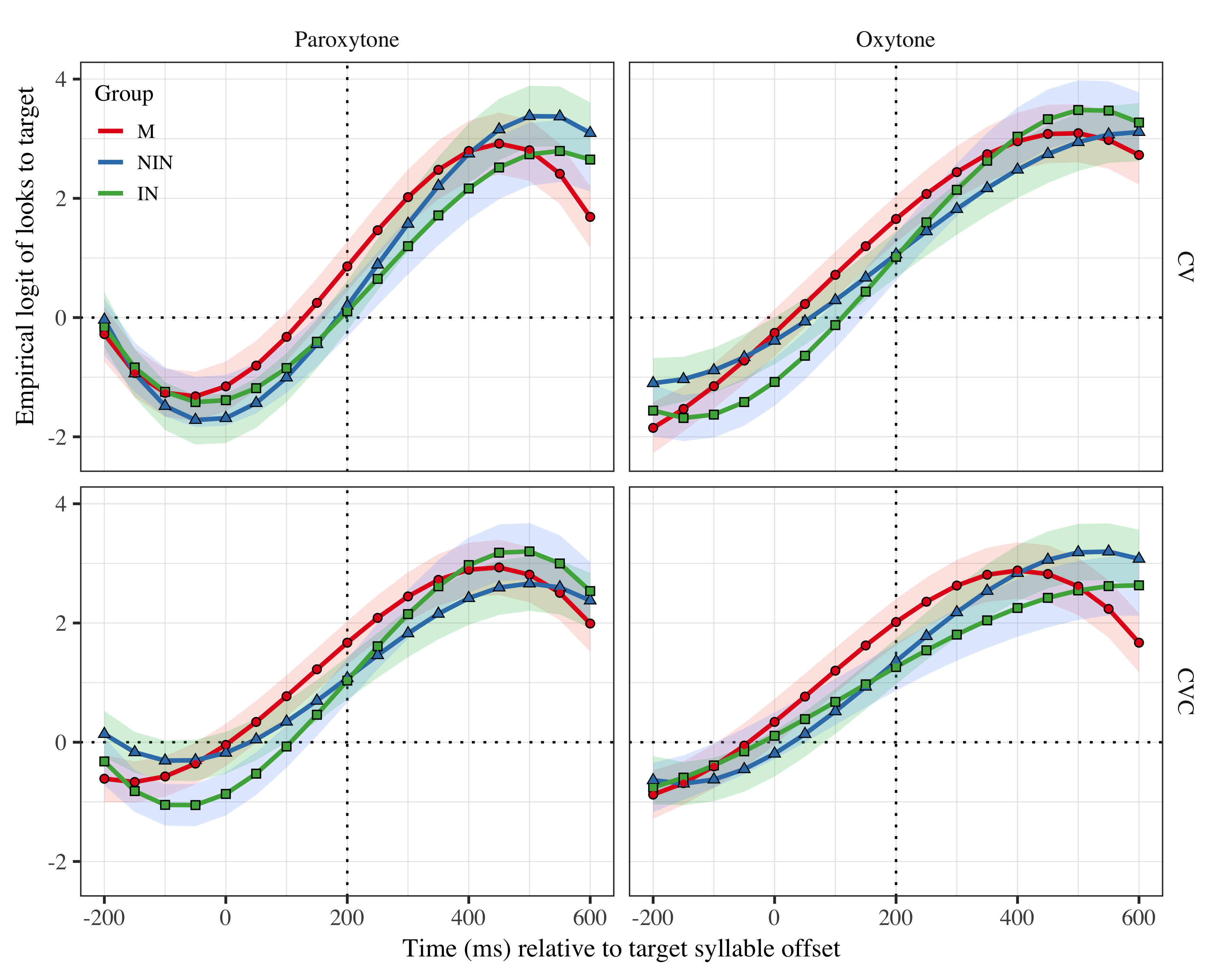


Figure 2: Growth curve estimates of target fixations as a function of lexical stress and syllable structure for each group during the analysis window. Symbols and lines represent model estimates, and the transparent ribbons represents ±SE. Empirical logit values on y-axis correspond to proportions of 0.12 0.50 0.88 0.98. The horizontal dotted line represents the 50% probability of fixating on the target. The vertical dotted line indicates 200 ms after the offset of the target syllable.

# Tables

| Group | Syllable structure | Lexical stress | Probability | LB | UB |
| --- | --- | --- | --- | --- | --- |
| M | CV | Paroxytone | 0.702 | 0.608 | 0.782 |
|  |  | Oxytone | 0.839 | 0.779 | 0.886 |
|  | CVC | Paroxytone | 0.842 | 0.787 | 0.884 |
|  |  | Oxytone | 0.882 | 0.836 | 0.917 |
| NIN | CV | Paroxytone | 0.550 | 0.446 | 0.649 |
|  |  | Oxytone | 0.742 | 0.661 | 0.810 |
|  | CVC | Paroxytone | 0.745 | 0.672 | 0.807 |
|  |  | Oxytone | 0.795 | 0.726 | 0.851 |
| IN | CV | Paroxytone | 0.526 | 0.420 | 0.629 |
|  |  | Oxytone | 0.735 | 0.650 | 0.805 |
|  | CVC | Paroxytone | 0.738 | 0.661 | 0.802 |
|  |  | Oxytone | 0.779 | 0.704 | 0.840 |

Table 1: Probability of target fixations ±SE at 200 ms after the target syllable offset.

## Fixed effects

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | Estimate | SE | *t* | *p* |
| Intercept (γ00) | 1.167 | 0.306 | 3.810 | < .001 |
| Time1 (γ10) | 5.704 | 1.042 | 5.476 | < .001 |
| Time2 (γ20) | −1.373 | 0.423 | −3.246 | .001 |
| Time3 (γ30) | −1.711 | 0.367 | −4.658 | < .001 |
| Syllable structure (γ01) | −0.074 | 0.203 | −0.365 | .715 |
| Time1 × Syllable structure (γ11) | 0.772 | 0.621 | 1.243 | .214 |
| Time2 × Syllable structure (γ21) | 0.571 | 0.310 | 1.842 | .066 |
| Time3 × Syllable structure (γ31) | −0.594 | 0.260 | −2.283 | .022 |
| Lexical stress (γ02) | −0.092 | 0.246 | −0.373 | .709 |
| Time1 × Lexical stress (γ12) | 0.125 | 0.616 | 0.203 | .839 |
| Time2 × Lexical stress (γ22) | 0.666 | 0.305 | 2.184 | .029 |
| Time3 × Lexical stress (γ32) | −0.325 | 0.256 | −1.269 | .204 |
| Group NIN (γ03) | −0.131 | 0.277 | −0.472 | .637 |
| Time1 × Group NIN (γ13) | 0.365 | 0.912 | 0.401 | .689 |
| Time2 × Group NIN (γ23) | 1.819 | 0.448 | 4.060 | < .001 |
| Time3 × Group NIN (γ33) | 0.124 | 0.385 | 0.323 | .747 |
| Group IN (γ04) | −0.255 | 0.287 | −0.889 | .374 |
| Time1 × Group IN (γ14) | 0.668 | 0.942 | 0.709 | .478 |
| Time2 × Group IN (γ24) | 1.615 | 0.462 | 3.496 | < .001 |
| Time3 × Group IN (γ34) | 0.022 | 0.396 | 0.056 | .956 |
| Syllable structure × Lexical stress (γ05) | −0.029 | 0.126 | −0.233 | .816 |
| Time1 × Syllable structure × Lexical stress (γ15) | −1.047 | 0.464 | −2.255 | .024 |
| Time2 × Syllable structure × Lexical stress (γ25) | 0.146 | 0.282 | 0.517 | .605 |
| Time3 × Syllable structure × Lexical stress (γ35) | −0.405 | 0.224 | −1.811 | .070 |
| Syllable structure × Lexical stress × Group NIN (γ06) | 0.028 | 0.067 | 0.425 | .671 |
| Time1 × Syllable structure × Lexical stress × Group NIN (γ16) | 1.004 | 0.271 | 3.708 | < .001 |
| Time2 × Syllable structure × Lexical stress × Group NIN (γ26) | 0.219 | 0.269 | 0.815 | .415 |
| Time3 × Syllable structure × Lexical stress × Group NIN (γ36) | −0.034 | 0.267 | −0.127 | .899 |
| Syllable structure × Lexical stress × Group IN (γ07) | −0.014 | 0.069 | −0.199 | .842 |
| Time1 × Syllable structure × Lexical stress × Group IN (γ17) | −0.507 | 0.278 | −1.821 | .069 |
| Time2 × Syllable structure × Lexical stress × Group IN (γ27) | 0.166 | 0.277 | 0.600 | .548 |
| Time3 × Syllable structure × Lexical stress × Group IN (γ37) | 0.773 | 0.275 | 2.816 | .005 |

Table 2: Growth curve model fixed effects (must reduce font size and single space).

## Random effects

Table 1: Model random effects

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Group | Parameter | Variance | SD | Correlations |  |  |  |  |  |
| participant | Intercept | 0.911 | 0.954 | 1.00 |  |  |  |  |  |
|  | Syllable structure | 0.275 | 0.524 | −.20 | 1.00 |  |  |  |  |
|  | Lexical stress | 0.789 | 0.888 | −.07 | .31 | 1.00 |  |  |  |
|  | Time1 | 9.548 | 3.090 | .42 | −.17 | .02 | 1.00 |  |  |
|  | Time2 | 1.640 | 1.281 | −.14 | .22 | .08 | .31 | 1.00 |  |
|  | Time3 | 0.980 | 0.990 | −.40 | .08 | −.18 | −.83 | −.14 | 1.00 |
| target | Intercept | 0.264 | 0.514 | 1.00 |  |  |  |  |  |
|  | Time1 | 3.831 | 1.957 | .28 |  |  | 1.00 |  |  |
|  | Time2 | 1.304 | 1.142 | −.74 |  |  | −.37 | 1.00 |  |
|  | Time3 | 0.415 | 0.644 | .19 |  |  | −.86 | −.14 | 1.00 |
| Residual |  | 13.507 | 3.675 |  |  |  |  |  |  |

# Statistical Analyses

The time course data from the eye-tracking task were analized using growth curve analysis (GCA, Mirman, 2016). We downsampled the data to bins of 50 ms which were centered at the offset of the first syllable of target items. The time course of fixation ranged from 200 ms before target syllable offset to 600 ms after. The empirical logit transformation (Barr, 2008) was applied to the binary responses (fixations to the target or the distractor). We modeled the time course using linear, quadratic, and cubic orthogonal polynomials with fixed effects of group, lexical stress, and syllable structure on all time terms. For the group predictor M was set as the baseline, thus the IN and NIN parameters described how the growth curve of the learners differed from that of the native controls. Lexical stress and syllable structure were sum coded such that parameter estimates represent effect sizes of change from CV to CVC syllables and paroxytone to oxytone stress. All models included by-subject random effects on all time terms and the syllable structure and lexical stress predictors, as well as by-item random effects on all time terms. Main effects and higher order interactions were assessed using nested model comparisons. The analysis was conducted in R (R Core Team, 2019) and the GCA models were fit using lme4 (Bates, Mächler, Bolker, & Walker, 2009). Pairwise comparisons between learners groups were conducted using the R package multcomp (Hothorn, Bretz, & Westfall, 2008).

# Results

The was a main effect of something on something (*χ*2(2) = 11, *p* = .004).

The effect was awesome (γ00 = 1.18; SE = 0.22; *t* = 5.36; *p* < .001).

build individual models for each group - model intercept tests if they predict above chance over time course - understand coda x condition interaction for each group build big model to test group differences - examine how learners differ from natives - pairwise tests to see how learners differ from each other - these 4-way interactions are difficult to understand - we can examine stress and syllable structure effect size over time course

ind\_mods$gca\_mod\_ss\_int\_3 %>% tidy\_lme4 %>% mutate(p = format\_pval(p))

gca\_mod\_ss\_int\_1 %>% tidy\_lme4 %>% mutate(p = format\_pval(p))