Contact helps dispreferred combinations of typological features to survive: geospatial evidence

Statistical analyses

Preliminaries

```
# required packages
require(tidyverse)
```

```
## Loading required package: tidyverse
```

```
## -- Attaching packages --------------------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.1      v purrr   1.0.1
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.3.0      v stringr 1.5.0
## v readr   2.1.4      v forcats 1.0.0
## -- Conflicts ------------------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
require(broom)
```

```
## Loading required package: broom
```

```
require(pixiedust)
```

```
## Loading required package: pixiedust
```

```
require(emmeans)
```

```
## Loading required package: emmeans
```

```
# create results/tables, if it doesn't exist already
try(dir.create("../results/tables", recursive=TRUE))
```

```
## Warning in dir.create("../results/tables", recursive = TRUE):
## '../results/tables' already exists
```

```
# load data
source("load_data.R")
```

```
## [1] 12
## [1] 14
```

Basic model: comparison of Δ between typologies of different statuses

Under-represented types (Δ^-)

```
mod_w <- lm(Delta_under ~ status+abs(phi), data=wals)
mod_g <- lm(Delta_under ~ status+abs(phi), data=gram)

#mod_w %>% dust %>% sprinkle(round=5) %>% write.csv(file="../results/tables/mod1_under_wals.csv", row.n
#mod_g %>% dust %>% sprinkle(round=5) %>% write.csv(file="../results/tables/mod1_under_grambank.csv", r

print(summary(mod_w))
```

```
##
## Call:
## lm(formula = Delta_under ~ status + abs(phi), data = wals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.16764 -0.05691 -0.02294  0.04428  0.28278
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.11470    0.02934   3.910 0.000216 ***
## statusinteracting  0.09716    0.04213   2.306 0.024150 *
## statusunknown      0.03630    0.03065   1.184 0.240414
## abs(phi)          -0.10699    0.07028  -1.522 0.132556
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1007 on 68 degrees of freedom
## Multiple R-squared:  0.07276,    Adjusted R-squared:  0.03185
## F-statistic: 1.779 on 3 and 68 DF,  p-value: 0.1595
```

```
print(summary(mod_g))
```

```
##
## Call:
## lm(formula = Delta_under ~ status + abs(phi), data = gram)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.283602 -0.077034 -0.000193  0.096815  0.265371
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.14587    0.03087   4.725  1.2e-05 ***
## statusinteracting  0.16956    0.05707   2.971   0.0041 **
## statusunknown      0.08898    0.03802   2.340   0.0222 *
## abs(phi)          -0.13914    0.07024  -1.981   0.0516 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1227 on 68 degrees of freedom
```

```
## Multiple R-squared:  0.1233, Adjusted R-squared:  0.08461
## F-statistic: 3.188 on 3 and 68 DF,  p-value: 0.02913
```

Over-represented types (Δ^+)

```
mod_w <- lm(Delta_over ~ status+abs(phi), data=wals)
mod_g <- lm(Delta_over ~ status+abs(phi), data=gram)

#mod_w %>% dust %>% sprinkle(round=5) %>% write.csv(file="../results/tables/mod1_over_wals.csv", row.na
#mod_g %>% dust %>% sprinkle(round=5) %>% write.csv(file="../results/tables/mod1_over_grambank.csv", ro

print(summary(mod_w))
```

```
##
## Call:
## lm(formula = Delta_over ~ status + abs(phi), data = wals)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.189314 -0.003346  0.014428  0.027497  0.051457
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -0.03357    0.01343  -2.500   0.0148 *
## statusinteracting -0.02346    0.01928  -1.217   0.2280
## statusunknown     -0.03318    0.01403  -2.365   0.0209 *
## abs(phi)           0.03125    0.03217   0.972   0.3347
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04608 on 68 degrees of freedom
## Multiple R-squared:  0.09165,    Adjusted R-squared:  0.05158
## F-statistic: 2.287 on 3 and 68 DF,  p-value: 0.0864
```

```
print(summary(mod_g))
```

```
##
## Call:
## lm(formula = Delta_over ~ status + abs(phi), data = gram)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.147464 -0.012747  0.007731  0.039623  0.073455
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -0.04436    0.01289  -3.442 0.000991 ***
## statusinteracting -0.04219    0.02383  -1.771 0.081082 .
## statusunknown     -0.02917    0.01587  -1.837 0.070507 .
## abs(phi)           0.01959    0.02932   0.668 0.506465
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05124 on 68 degrees of freedom
## Multiple R-squared:  0.05754,    Adjusted R-squared:  0.01596
```

```
## F-statistic: 1.384 on 3 and 68 DF,  p-value: 0.2551
```

Model 2: model comparison between φ and φ_c as predictors

Under-represented types

```
mod_w <- lm(Delta_under ~ abs(phi), data=wals)
mod_wc <- lm(Delta_under ~ abs(corrected_phi), data=wals)

print(summary(mod_w))
```

```
##
## Call:
## lm(formula = Delta_under ~ abs(phi), data = wals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.14322 -0.06548 -0.02422  0.06118  0.26854
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.125997   0.021852   5.766 2.02e-07 ***
## abs(phi)    -0.004628   0.052343  -0.088     0.93
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.103 on 70 degrees of freedom
## Multiple R-squared:  0.0001117,  Adjusted R-squared:  -0.01417
## F-statistic: 0.007819 on 1 and 70 DF,  p-value: 0.9298
```

```
print(summary(mod_wc))
```

```
##
## Call:
## lm(formula = Delta_under ~ abs(corrected_phi), data = wals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.12985 -0.06951 -0.02189  0.03172  0.29817
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.09490    0.01898   5.000 4.07e-06 ***
## abs(corrected_phi)  0.13612    0.06855   1.986    0.051 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1003 on 70 degrees of freedom
## Multiple R-squared:  0.05333,    Adjusted R-squared:  0.0398
## F-statistic: 3.943 on 1 and 70 DF,  p-value: 0.05098
```

```
print(AIC(mod_w))
```

```
## [1] -118.9743
```

```
print(AIC(mod_wc))
```

```
## [1] -122.912
```

```
mod_g <- lm(Delta_under ~ abs(phi), data=gram)
mod_gc <- lm(Delta_under ~ abs(corrected_phi), data=gram)

print(summary(mod_g))
```

```
##
## Call:
## lm(formula = Delta_under ~ abs(phi), data = gram)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.25096 -0.09167 -0.02315  0.10416  0.29348
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.20220    0.02093   9.662 1.63e-14 ***
## abs(phi)    -0.03236    0.06383  -0.507    0.614
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.129 on 70 degrees of freedom
## Multiple R-squared:  0.003658,   Adjusted R-squared:  -0.01058
## F-statistic: 0.257 on 1 and 70 DF,  p-value: 0.6138
```

```
print(summary(mod_gc))
```

```
##
## Call:
## lm(formula = Delta_under ~ abs(corrected_phi), data = gram)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.24447 -0.08134 -0.02812  0.10035  0.31496
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.17389    0.02064   8.424 3.03e-12 ***
## abs(corrected_phi)  0.13634    0.09204   1.481    0.143
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1272 on 70 degrees of freedom
## Multiple R-squared:  0.03039,    Adjusted R-squared:  0.01654
## F-statistic: 2.194 on 1 and 70 DF,  p-value: 0.143
```

```
print(AIC(mod_g))
```

```
## [1] -86.65324
```

```
print(AIC(mod_gc))
```

```
## [1] -88.61157
```

```
#mod_w %>% dust %>% sprinkle(round=5) %>% write.csv(file="../results/tables/mod2_under_wals.csv", row.n
#mod_g %>% dust %>% sprinkle(round=5) %>% write.csv(file="../results/tables/mod2_under_grambank.csv", r
#mod_wc %>% dust %>% sprinkle(round=5) %>% write.csv(file="../results/tables/mod2_under_corrected_wals.
#mod_gc %>% dust %>% sprinkle(round=5) %>% write.csv(file="../results/tables/mod2_under_corrected_gramb
```

Over-represented types

```
mod_w <- lm(Delta_over ~ abs(phi), data=wals)
mod_wc <- lm(Delta_over ~ abs(corrected_phi), data=wals)

print(summary(mod_w))
```

```
##
## Call:
## lm(formula = Delta_over ~ abs(phi), data = wals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.18919 -0.01275  0.01277  0.02911  0.05472
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.05467    0.01002  -5.456 6.95e-07 ***
## abs(phi)     0.02626    0.02400   1.094    0.278
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04725 on 70 degrees of freedom
## Multiple R-squared:  0.01681,    Adjusted R-squared:  0.002769
## F-statistic: 1.197 on 1 and 70 DF,  p-value: 0.2776
```

```
print(summary(mod_wc))
```

```
##
## Call:
## lm(formula = Delta_over ~ abs(corrected_phi), data = wals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.18662 -0.01992  0.01509  0.03255  0.04789
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -0.047468   0.009016  -5.265 1.47e-06 ***
## abs(corrected_phi)  0.008834   0.032563   0.271    0.787
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04762 on 70 degrees of freedom
## Multiple R-squared:  0.00105,    Adjusted R-squared:  -0.01322
## F-statistic: 0.07359 on 1 and 70 DF,  p-value: 0.787
```

```
print(AIC(mod_w))
```

```
## [1] -231.2405
```

```
print(AIC(mod_wc))
```

```
## [1] -230.0953
```

```
mod_g <- lm(Delta_over ~ abs(phi), data=gram)
mod_gc <- lm(Delta_over ~ abs(corrected_phi), data=gram)

print(summary(mod_g))
```

```
##
## Call:
## lm(formula = Delta_over ~ abs(phi), data = gram)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.157168 -0.020617  0.004217  0.041668  0.064181
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.062654   0.008436  -7.427 2.07e-10 ***
## abs(phi)    -0.007939   0.025732  -0.309    0.759
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05199 on 70 degrees of freedom
## Multiple R-squared:  0.001358,   Adjusted R-squared:  -0.01291
## F-statistic: 0.09519 on 1 and 70 DF,  p-value: 0.7586
```

```
print(summary(mod_gc))
```

```
##
## Call:
## lm(formula = Delta_over ~ abs(corrected_phi), data = gram)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.157237 -0.018750  0.002716  0.040443  0.065767
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -0.062390   0.008434  -7.398 2.34e-10 ***
## abs(corrected_phi) -0.013320   0.037604  -0.354    0.724
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05198 on 70 degrees of freedom
## Multiple R-squared:  0.001789,   Adjusted R-squared:  -0.01247
## F-statistic: 0.1255 on 1 and 70 DF,  p-value: 0.7243
```

```
print(AIC(mod_g))
```

```
## [1] -217.4769
```

```
print(AIC(mod_gc))
```

```
## [1] -217.508
```