

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

Statistical analyses

Preliminaries

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

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

```
## -- Attaching core tidyverse packages -----
```

```
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr    1.5.1
## v ggplot2    3.5.1      v tibble     3.2.1
## v lubridate  1.9.3      v tidyr      1.3.1
## v purrr      1.0.2
```

```
## -- Conflicts -----
```

```
## x dplyr::filter() masks stats::filter()
```

```
## x dplyr::lag()     masks stats::lag()
```

```
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
require(broom)
```

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

```
require(pixiedust)
```

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

```
require(emmeans)
```

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

```
## Welcome to emmeans.
```

```
## Caution: You lose important information if you filter this package's results.
```

```
## See '? untidy'
```

```
# 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] 27
```

```
## [1] 36
```

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.names=FALSE)
#mod_g %>% dust %>% sprinkle(round=5) %>% write.csv(file="../results/tables/mod1_under_grambank.csv", row.names=FALSE)

print(summary(mod_w))
```

```
##
## Call:
## lm(formula = Delta_under ~ status + abs(phi), data = wals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.07117 -0.03676 -0.01405  0.02890  0.11073
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.09566    0.03002   3.187  0.00279 **
## statusunknown    -0.03212    0.02379  -1.350  0.18459
## statusnon-interacting -0.05300    0.02513  -2.109  0.04127 *
## abs(phi)         -0.05168    0.04488  -1.152  0.25635
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04804 on 40 degrees of freedom
## Multiple R-squared:  0.1045, Adjusted R-squared:  0.03737
## F-statistic: 1.556 on 3 and 40 DF,  p-value: 0.2149
```

```
print(summary(mod_g))

##
## Call:
## lm(formula = Delta_under ~ status + abs(phi), data = gram)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.087397 -0.040869  0.000164  0.033713  0.112064
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.13399    0.03151   4.253  0.000123 ***
## statusunknown    -0.03904    0.02854  -1.368  0.178951
## statusnon-interacting -0.09311    0.03314  -2.809  0.007641 **
## abs(phi)         -0.05644    0.04398  -1.283  0.206772
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05454 on 40 degrees of freedom
```

```
## Multiple R-squared:  0.2093, Adjusted R-squared:  0.15
## F-statistic:  3.53 on 3 and 40 DF,  p-value: 0.02324
```

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.names=FALSE)
#mod_g %>% dust %>% sprinkle(round=5) %>% write.csv(file="../results/tables/mod1_over_grambank.csv", row.names=FALSE)

print(summary(mod_w))
```

```
##
## Call:
## lm(formula = Delta_over ~ status + abs(phi), data = wals)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-0.069612	-0.000688	0.007237	0.011959	0.022515

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.009284	0.013376	-0.694	0.492
statusunknown	-0.010437	0.010601	-0.985	0.331
statusnon-interacting	0.003184	0.011200	0.284	0.778
abs(phi)	-0.006369	0.019999	-0.318	0.752

```
##
## Residual standard error: 0.02141 on 40 degrees of freedom
## Multiple R-squared:  0.08283,    Adjusted R-squared:  0.01405
## F-statistic: 1.204 on 3 and 40 DF,  p-value: 0.3206
```

```
print(summary(mod_g))

##
## Call:
## lm(formula = Delta_over ~ status + abs(phi), data = gram)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-0.040862	-0.005200	0.005317	0.010354	0.018148

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.0064249	0.0087067	-0.738	0.465
statusunknown	-0.0040992	0.0078874	-0.520	0.606
statusnon-interacting	0.0009538	0.0091589	0.104	0.918
abs(phi)	-0.0143397	0.0121546	-1.180	0.245

```
##
## Residual standard error: 0.01507 on 40 degrees of freedom
## Multiple R-squared:  0.09068,    Adjusted R-squared:  0.02248
## F-statistic:  1.33 on 3 and 40 DF,  p-value: 0.2782
```

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
## -0.05838 -0.03749 -0.01883  0.02313  0.11344
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.044783   0.013170   3.400  0.00149 **
## abs(phi)      0.007748   0.033594   0.231  0.81871
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04951 on 42 degrees of freedom
## Multiple R-squared:  0.001265, Adjusted R-squared:  -0.02251
## F-statistic: 0.0532 on 1 and 42 DF, p-value: 0.8187

print(summary(mod_wc))

##
## Call:
## lm(formula = Delta_under ~ abs(corrected_phi), data = wals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.06164 -0.03415 -0.01189  0.02703  0.11328
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.02732   0.01072   2.548  0.0146 *
## abs(corrected_phi) 0.10336   0.04209   2.455  0.0183 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04633 on 42 degrees of freedom
## Multiple R-squared:  0.1255, Adjusted R-squared:  0.1047
## F-statistic: 6.029 on 1 and 42 DF, p-value: 0.01829

print(AIC(mod_w))

## [1] -135.6677

print(AIC(mod_wc))

## [1] -141.5141
```

```
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
	-0.08047	-0.04808	-0.01700	0.04190	0.13964

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.06465	0.01182	5.470	2.29e-06 ***
abs(phi)	0.02159	0.03970	0.544	0.59

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05965 on 42 degrees of freedom
## Multiple R-squared:  0.006988, Adjusted R-squared:  -0.01666
## F-statistic: 0.2956 on 1 and 42 DF, p-value: 0.5896
```

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

```
##
## Call:
## lm(formula = Delta_under ~ abs(corrected_phi), data = gram)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-0.07302	-0.04413	-0.02022	0.02559	0.14432

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.05496	0.01205	4.559	4.39e-05 ***
abs(corrected_phi)	0.09604	0.05753	1.669	0.102

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05797 on 42 degrees of freedom
## Multiple R-squared:  0.06223, Adjusted R-squared:  0.0399
## F-statistic: 2.787 on 1 and 42 DF, p-value: 0.1025
```

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

```
## [1] -119.2699
```

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

```
## [1] -121.7884
```

```
#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
## -0.077625 -0.003407  0.008785  0.014097  0.016114
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.013129   0.005799  -2.264   0.0288 *
## abs(phi)     -0.003363   0.014792  -0.227   0.8212
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0218 on 42 degrees of freedom
## Multiple R-squared:  0.001229, Adjusted R-squared:  -0.02255
## F-statistic: 0.0517 on 1 and 42 DF, p-value: 0.8212

print(summary(mod_wc))

##
## Call:
## lm(formula = Delta_over ~ abs(corrected_phi), data = wals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.078422 -0.003175  0.009555  0.013748  0.016292
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -0.013318   0.005044  -2.640   0.0116 *
## abs(corrected_phi) -0.004647   0.019807  -0.235   0.8157
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0218 on 42 degrees of freedom
## Multiple R-squared:  0.001309, Adjusted R-squared:  -0.02247
## F-statistic: 0.05504 on 1 and 42 DF, p-value: 0.8157

print(AIC(mod_w))

## [1] -207.8521

print(AIC(mod_wc))
```

```
## [1] -207.8556
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
## -0.042483 -0.003295  0.007537  0.008966  0.016760
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.007675   0.002953  -2.599   0.0128 *
## abs(phi)     -0.017167   0.009920  -1.731   0.0909 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0149 on 42 degrees of freedom
## Multiple R-squared:  0.06656,    Adjusted R-squared:  0.04433
## F-statistic: 2.995 on 1 and 42 DF,  p-value: 0.09087

print(summary(mod_gc))

##
## Call:
## lm(formula = Delta_over ~ abs(corrected_phi), data = gram)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.046902 -0.003097  0.003287  0.010824  0.011298
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -0.010697   0.003207  -3.335   0.00179 **
## abs(corrected_phi) -0.002037   0.015305  -0.133   0.89478
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01542 on 42 degrees of freedom
## Multiple R-squared:  0.0004214,    Adjusted R-squared:  -0.02338
## F-statistic: 0.01771 on 1 and 42 DF,  p-value: 0.8948

print(AIC(mod_g))

## [1] -241.3176

print(AIC(mod_gc))

## [1] -238.3055
```