```
library(MASS)
library(lme4)
library(lattice)
library(ez)
library(multcomp)
source(file.path(pfadu, "phoc.txt"))
source(file.path(pfadu, "sigmoid.txt"))

# 1.
# 10 Sprecher produzierten unbetonte und betonte Silben.
# Die Dauern ihrer Silben waren wie folgt:

# unbetont
un = c(165, 178, 143, 187, 186, 127, 138, 155, 157, 171)

# betont
be = c(182, 189, 179, 196, 188, 153, 154, 178, 169, 191)

# Hat die Betonung einen Einfluss auf die Dauer?
boxplot(un - be)
t.test(un  - be)
# Die Dauern wurden signifikant von der Betonung beeinflusst
# (t[9] = 5.6, p < 0.001)
# oder

werte = c(un, be)
werte.l =  c(rep("un", length(un)), rep("be", length(be)))
vpn = rep(paste("S", 1:10, sep=""), 2)
w = data.frame(Vpn = factor(vpn), werte, Be = factor(werte.l))
t.test(werte ~ Be, paired=T, data = w)
# oder
ezANOVA(w, .(werte), .(Vpn), .(Be))

# 2.
# Die dB-Werte in der Lösung von einem /t/ für deutsche und
# französische Sprecher waren wie folgt. Hat die Sprachgruppe einen Einfluss auf den
dB-Wert?
# Deutsch
deutsch = c(68, 98, 99, 88, 78, 91, 69, 91, 83, 83, 74, 57, 91, 91, 97, 87, 90, 84,
73, 53)

# Französisch
franz = c(80, 60, 82, 58, 49, 52, 47, 92, 76, 79, 88, 89, 51, 72, 95, 62, 75, 63, 59)
db = c(deutsch, franz)
db.l = c(rep("deutsch", length(deutsch)), rep("franz", length(franz)))
vpn = paste("S", 1:length(db.l))
loesung = data.frame(db, Sprache = factor(db.l), Vpn = factor(vpn))
boxplot(db ~ Sprache, data = loesung)
t.test(db ~ Sprache, data = loesung)
# Die dB-Werte wurden signifikant von der Sprache beeinflusst
# (t[35.2] = 2.7, p < 0.05)

# 3.
```

```
# 10 Sprecher produzierten Wörter in der Laborsprache, der Lesesprache, und
# in der Spontansprache. Die jeweiligen Silbendauern dieser Sprecher sind wie folgt:

# Spontan
spont = c(12, 18, 15, 21, 19, 10, 18, 19, 23, 17)

# Lesesprache
lese = c(15, 21, 16, 26, 23, 12, 17, 17, 27, 25)

# Laborsprache
labor = c(18, 19, 15, 32, 22, 14, 21, 21, 30, 21)

# Hat der Sprechstil einen Einfluss auf die Silbendauer?

d = c(spont, lese, labor)
lab = c(rep("spont", length(spont)), rep("lese", length(lese)), rep("labor", length
(labor)))
vpn = rep(paste("S", 1:10, sep=""), 3)
stil = data.frame(d, Stil = factor(lab), Vpn = factor(vpn))
boxplot(d ~ Stil, data = stil)
# labor > lese > spont
ezANOVA(stil, .(d), .(Vpn), .(Stil))
 phoc(stil, .(d), .(Vpn), .(Stil))

# NB: GGe > 0.75, daher HFe, aber da HFe > 1, kann Sphericity ignoriert werden
# Die Dauer wurde signifikant vom Sprechstil beeinflusst (F[2,18] = 9.0, p < 0.01).
# Post-hoc Tests zeigten signifikante Unterschiede zwischen
# der Spontan- und Lesesprache (p < 0.05) und zwischen der
# Spontansprache und Laborsprache (p < 0.01) aber
# nicht zwischen der Lese- und Laborsprache


# 4.
dez = read.table(file.path(pfadu, "dez.txt"))
# Die Daten zeigen dB-Werte für /a/ Vokale produziert von 27 verschiedenen Sprechern
# aus drei Regionen (Faktor Region) und aufgeteilt in drei sozio-ökonomische
# Gruppen (Faktor SEG). Wurden die dB-Werte von der Region und/oder
# der sozioökonomischen Gruppe beeinflusst?
boxplot(db ~ Region * SEG, data = dez)
# A > B > C; R1 > R2 > R3
ezANOVA(dez, .(db), .(Vpn), between = .(Region, SEG))
# Die dB-Werte wurden signifikant von
# der Region (F[2, 18] = 64.1, p < 0.001) und
# von SEG (F[2,18] = 40.2, p < 0.001) beeinflusst
# und es gab keine Interaktion zwischen diesen Faktoren.


# 5.
g = read.table(file.path(pfadu, "gfreq.txt"))
# Diese Daten zeigen die Grundfrequenzwerte für verschiedene Jahrgänge in derselben
Person.
# Prüfen Sie, ob ein lineares Verhältnis zwischen diesen Variablen besteht. Was wäre
# der vorhergesagte Grunfrequenzwert im 2015?
```

```
plot(Grund ~ Jahr, data = g)
# Linear und negativ
o = lm(Grund ~ Jahr, data = g)
abline(o)
predict(o, data.frame(Jahr = 2015))
# 88.6 Hz

# 6.
x = read.table(file.path(pfadu, "x24.txt"))

# Daten modifiziert aus Quené (http://www.let.uu.nl/~Hugo.Quene/personal/multilevel)
# Die Daten zeigen normalisierte
# Entfernungen von drei Vokalen (Faktor V) für verschiedene Wörter
# (Faktor (W)) und gesprochen von verschiedenen Sprechern (Faktor Vpn).
# Wurden die Entfernungen vom Vokal beeinflusst?

# Wir wollen hier Vpn und Wort als Random-Faktoren, da wir
# deren Variabilität ausklammern wollen (d.h. wir wollen
# messen, ob die Entfernungen vom Vokal beeinflusst wurden,
# unabhängig von der Variation, die wegen der Vpn oder Wort
# zustande kommt. Wir können einen boxplot machen, aber
# eventuell werden wir nicht viel sehen, wenn die Wort- und/oder
# Vpn-Variation groß ist
boxplot(Ent ~ V, data = x)
o = lmer(Ent ~ V + (1|Vpn) + (1|Wort), data = x)
ohne = update(o, ~ . -V)
anova(o, ohne)
summary(glht(o, linfct = mcp(V = "Tukey")))
# Die Entfernungen wurden signifikant von dem Vokal beeinflusst
# (c^2[2] = 45.9, p < 0.001). Post-hoc Tukey-Tests zeigten
# signifikante Unterschiede zwischen allen Vokal-Paaren
# (/i/ vs /a/: p < 0.001: /u/ vs /a/: p < 0.001; /u/ vs /i/: p < 0.05)

# Übrigens und nur zur Info sieht man manchmal mehr, wenn man z.B. über die Vpn
# mittelt und den Unterschied abbildet:
m = with(x, aggregate(Ent, list(Vpn, V), mean))
names(m) = c("Vpn", "V", "Ent")
# Mittelwert für /i/
temp = m$V == "i"
mi = m[temp,]
# Mittelwert für /a/
temp = m$V == "a"
ma = m[temp,]
# Mittelwert für /u/
temp = m$V == "u"
mu = m[temp,]
# i vs a Unterschied
boxplot(mi$Ent - ma$Ent)
# ist dies > 0?
t.test(mi$Ent - ma$Ent)
# i vs u Unterschied
boxplot(mi$Ent - mu$Ent)
# ist dies > 0? Nein - aber dies war auch p < 0.05 mit dem MM
```

```
t.test(mi$Ent - mu$Ent)
# a vs u Unterschied
boxplot(ma$Ent - mu$Ent)
# Ist dies < 0. Fast
t.test(ma$Ent - mu$Ent)
# In den letzten 2 Fällen ist das nicht ganz sig
# eventuell weil die Wort-Variabilität hoch ist (die wir
# in der obigen Berechnung nicht ausgeklammert haben)

# 7.
h = read.table(file.path(pfadu, "h24.txt"))
# Daten modifiziert aus Quené (http://www.let.uu.nl/~Hugo.Quene/personal/multilevel)
# Vokale wurden aus Wörtern (Faktor Wort) extrahiert und Versuchspersonen (Faktor Vpn
)
# präsentiert. Der Faktor Urteil zeigt, ob die Versuchspersonen die Vokale richtig (1
)
# oder falsch (0) identifiziert hatten. Wurde das Urteil von dem Vokal (Faktor V)
# beeinflusst?
tab = with(h, table(V, Urteil))
p = prop.table(tab, 1)
barchart(p, auto.key=T, horizontal=F)
# es könnte Unterschiede geben: vor allm /i/ - /a/ und /i/ - /u/
o = lmer(Urteil ~ V + (1|Wort) + (1|Vpn), family = binomial, data = h)
ohne = update(o, ~ . -V)
anova(o, ohne)
summary(glht(o, linfct = mcp(V = "Tukey")))
# Die Urteile wurden signifikant vom Vokal beeinflusst
# (c^[2] = 18.8, p < 0.001). Post-hoc Tukey Tests
# zeigten  einen signifikanten Unterschied
# zwischen /i/ und /a/ (p < 0.001); die Unterschiede
# zwischen /u/ und /a/ (p = 0.053) und zwischen
# /u/ und /i/ (p = 0.096) waren nicht ganz signifikant.

# 8.
japan = read.table(file.path(pfadu, "japan.txt"))
# (Daten von Yuki Era)
# Ein Kontinuum wurde erstellt, indem die Grundfrequenz in einem japanischen Satz in
11 Schritten
# herabgestuft wurde. Verschiedene Versuchspersonen mussten pro Stimulus
# beurteilen, ob der Satz eher wie eine Aussage (Aus) oder Erstauen (Ers) klingt
(Faktor Urteil).
# Berechnen Sie den Umkipppunkt für die Bevölkerung und überlagern Sie ihn auf
# die psychometrische Kurve zwischen Stimuli 1 und 11.
o = lmer(Urteil ~ Stim + (1+Stim|Vpn), family=binomial, data = japan)
k = fixef(o)[1]
m = fixef(o)[2]
tab = with(japan, table(Stim, Urteil))
p = prop.table(tab, 1)
sigmoid(p, k, m)
# Umkipppunkt überlagern
abline(v = -k/m)
```