

# Sprache ist im Kern ikonisch und multimodal

Berufungsvortrag für die Juniorprofessur für Phonetik (W1) mit Tenure Track  
(W2) an der Universität zu Köln

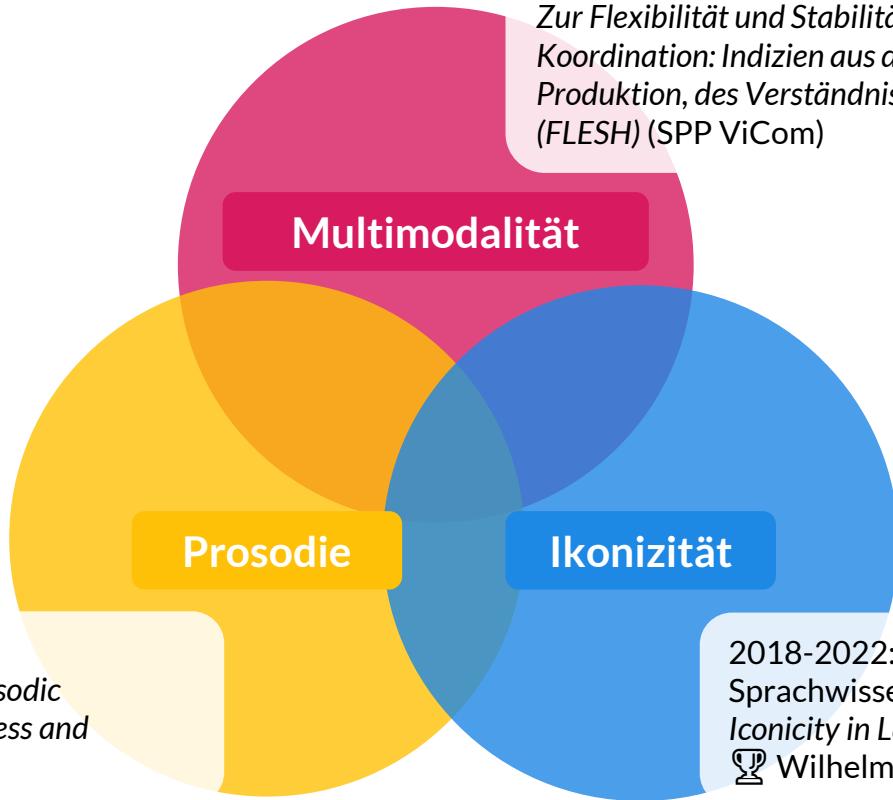
Aleksandra Ćwiek  
25. April 2024



Leibniz-Zentrum  
Allgemeine Sprachwissenschaft



2016-2018: M.A. Linguistik  
*The Acoustic Realization of Prosodic Prominence in Polish: Word Stress and Sentence Accent*

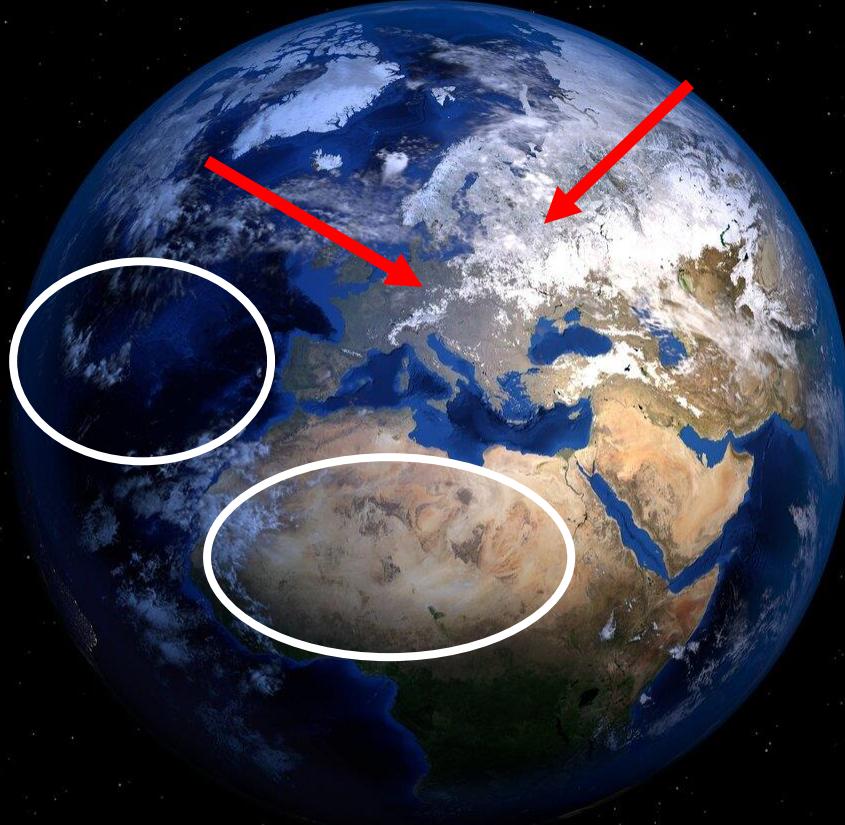


Seit 2022: **Projektleiterin** (ZAS)  
Zur Flexibilität und Stabilität der Gesten-Sprech-Koordination: Indizien aus der multimodalen Produktion, des Verständnisses und der Imitation (FLESH) (SPP ViCom)

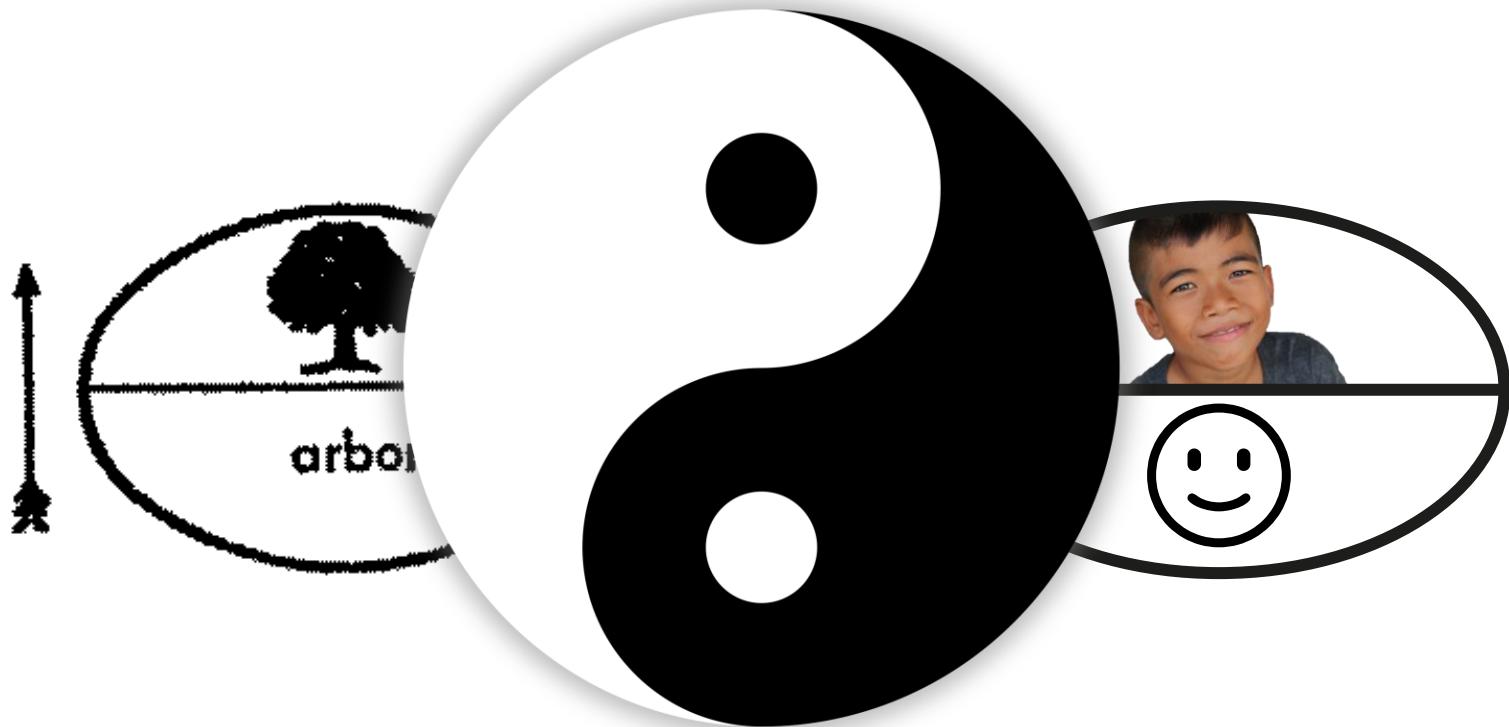
2018-2022: **Dr. phil.** Allgemeine Sprachwissenschaft (ZAS/HU Berlin)  
*Iconicity in Language and Speech*  
Wilhelm von Humboldt-Preis der DGfS

# *Language is iconic and multimodal to the core*

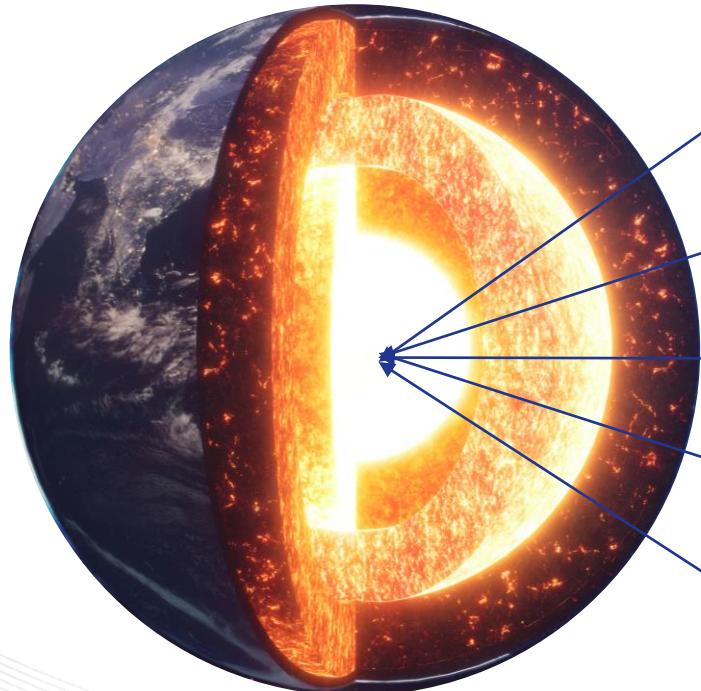
(Perlman, 2017)







De Saussure (1916); Perniss & Vigliocco (2014); Winter, Woodin & Perlman (2023)



### Phonologie (Lautsymbolik)

/a/ = groß, /i/ = klein  
(Sapir, 1929; Kilpatrick & Ćwiek, 2024)

### Prosodie

tief = groß, hoch = klein  
(Ohala, 1994; Ćwiek & Fuchs, 2019)

### Morpho(phono)logie

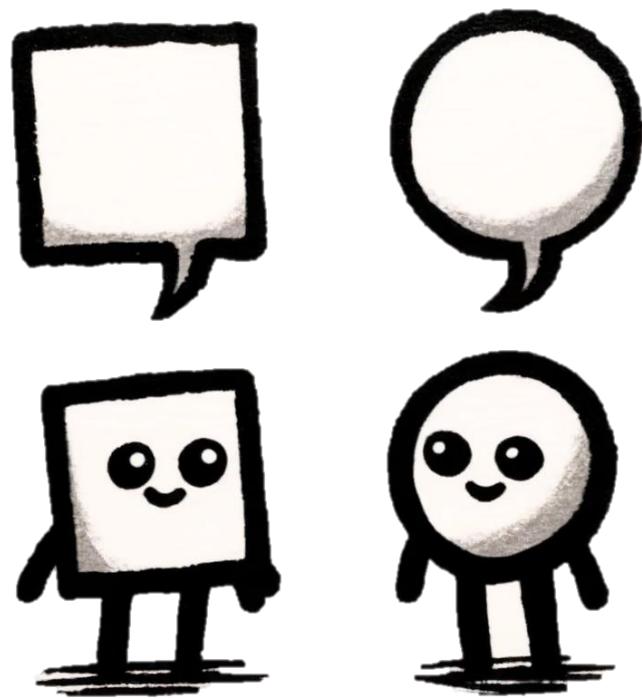
Reduplikation = Plural/Intensivierung  
(Himmelmann, 2005; Freywald, 2016)

### Morphosyntax/Lexikon (Ideophone)

ruckzuck, holterdipolter  
(Dingemanse, 2012; Ćwiek, 2022)

### Relevanz z.B. im Spracherwerb

(Imai & Kita, 2014; Perry et al., 2015)



## Verstehen ohne Worte: Der Ursprung der Sprache

Wenn uns im Ausland die Worte fehlen, dann versuchen wir es mit Gesten oder Lautmalereien. Und überrascht, dass wir verstanden werden. Die Gründe dafür liegen lange zurück, in der Entwicklung der Menschheit.

MDR KULTUR - Das Radio

Mi 12.05.2021 19:49 Uhr | 03:52 min

SEPTEMBER 1, 2021 | 2 MIN READ

## Made-Up Sounds Convey Meaning across Cultures

Newly created vocalizations can convey concepts remarkably well.

BY KATHERINE KORNEI



Am Anfang war die Lautmalerei

Berliner Forscher kommen zu dem Schluss, dass bei der Entstehung der Sprache offenbar nicht nur Gesten eine Rolle gespielt haben

Wortlose Verständigung

# Prägten Lautmalereien die Sprachevolution?

SPRACHENTSTEHUNG

## Lautmalereien werden überall verstanden

Am Beginn der menschlichen Sprachentstehung könnten Lautmalereien gestanden sein. Darauf deuten neue Experimente von Forscherinnen und Forschern hin. Sie haben dieselben Lautmalereien Menschen vorgespielt, die ganz unterschiedliche Sprachen sprechen – und wurden überwiegend verstanden.

NEWS | SOCIAL SCIENCES

## No shared language? No problem! People across cultures understand clues from 'vocal charades'

New study suggests how vocal sounds could help language get off the ground

14 MAY 2021 · BY CATHLEEN O'GRADY

## Sounds May Be Missing Link to Language

han gestures, could be key to birth of language

Tags: language language evolution e

Akustisch-auditive Ikonizität hat das  
Potenzial, ein **Kommunikationssystem**  
**zu gründen**

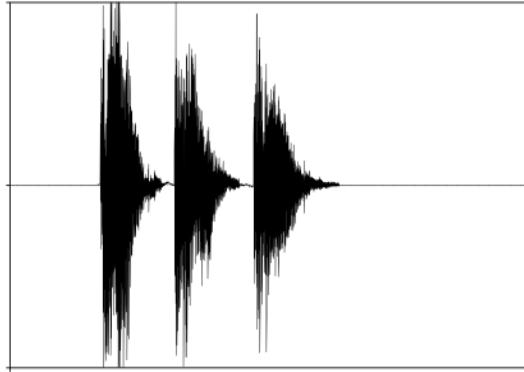


- 
- 01 ————— 02 ————— 03 ————— 04 →



Bildquelle: Herbert Terrace

Gardner & Gardner (1969); Hewes (1973); Hockett (1960, 1978); aber s. Fuchs & Ćwiek (2022); Kendon (2017)



verstecken

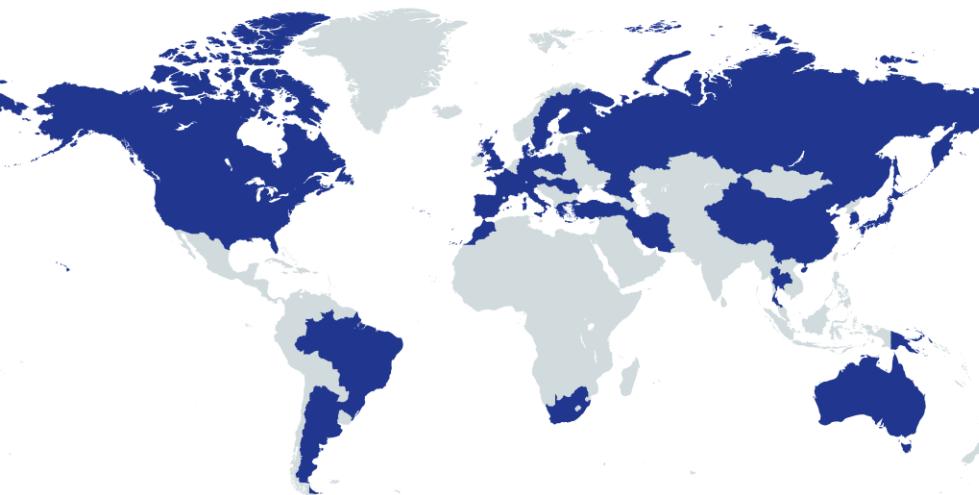
Mann

kochen

scharf

viele

jagen



## Online:

Albanisch	Italienisch	Russisch
Armenisch	Japanisch	Schwedisch
Deutsch	Koreanisch	Spanisch
Dänisch	Malaiisch	Tamilisch
Englisch (BE, AmE)	Malagasy	Thailändisch
Estnisch	Mandarin	Türkisch
Finnisch	Persisch	Ungarisch
Französisch	Polnisch	Zulu
Georgisch	Portugiesisch	
Griechisch	Rumänisch	

## Vor Ort:

Daakie  
Deutsch  
Englisch (BE, AmE)  
Palikúr  
Brasilianisches Portugiesisch  
Taschelhit-Berberisch

- Ikonische Vokalisierungen gehen weit über den Ausdruck von Gefühlen und die Nachahmung von Tieren hinaus:
  - Getestete Kategorien: *Handlungen, Menschen, Tiere, unbelebte Objekte, Eigenschaften, Quantifizierer, Demonstrativa*
- Hörer\*innen sind gut in der Lage, ikonische Vokalisierungen zu verstehen:
  - Online: 64% (Zufall: 16.7%)
  - Vor Ort: 55% (Zufall: 8%)

Ikonische Vokalisierungen sind ausdrucksstärker als gedacht!

Daten, Skripte, Modelle



**Das phonetische Potenzial der Ikonizität kann durch die Phonologie beeinflusst werden**

Akustisch-auditive Ikonizität hat das Potenzial, ein **Kommunikationssystem zu gründen**



**01**

**02**

**03**

**04**



**Welche Textur entspricht dieser Aufnahme?**

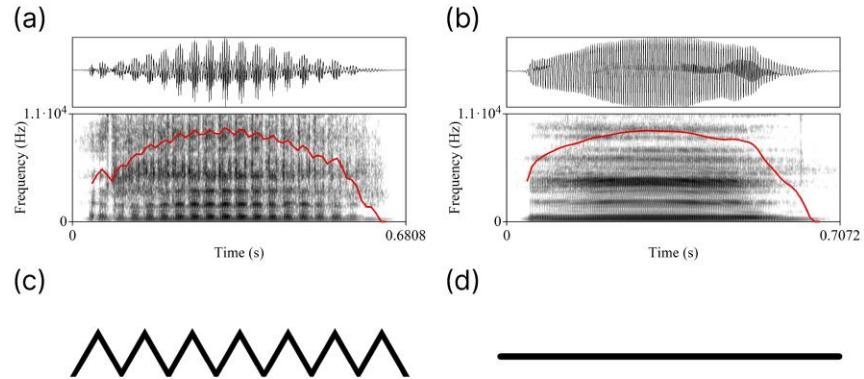
- [r] und [l] sind **Allophone eines Phonems** (kein Kontrast):
  - Online: Japanisch, Koreanisch, Mandarin, Zulu
  - Vor Ort: Palikúr
- [r] ist die **Hauptrealisierung** von /r/:
  - Online: Albanisch, Rumänisch, Polnisch, Armenisch, Spanisch, Persisch, Italienisch, Russisch, Estnisch, Ungarisch, Finnisch
  - Vor Ort: Daakie, Taschelhit-Berberisch

Daten, Skripte, Modelle

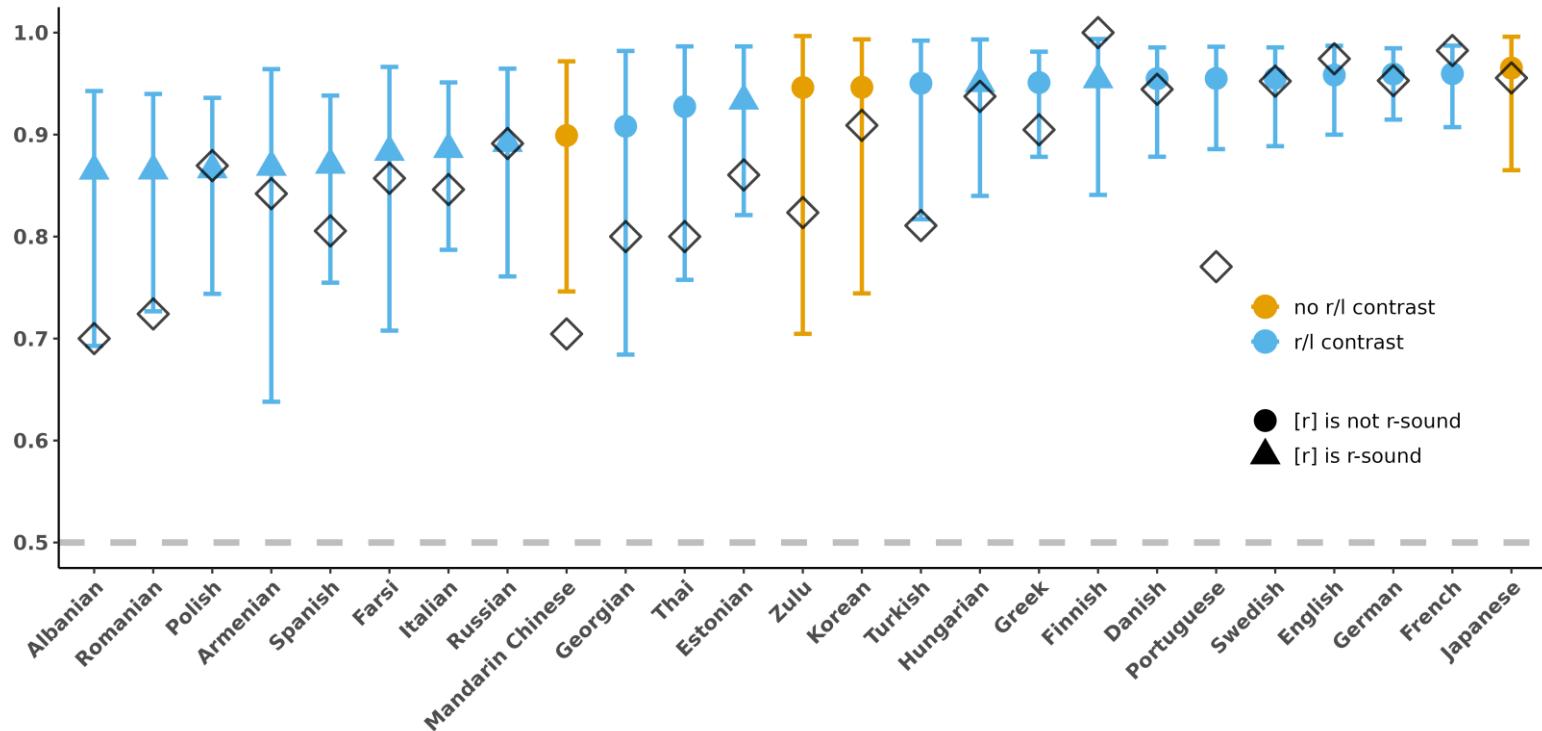


# Das Zusammenspiel der Phonetik und Phonologie in der cross-modalen Beziehung zwischen Klang und Tastsinn

- /r/ wird mit Rauheit verbunden (Winter et al., 2022)
- /r/ gilt als das prototypische r, aber hat nicht unbedingt etwas mit dem Trill zu tun (Anselme et al., 2022; 2023)
- Die cross-modale Verbindung zwischen /r/ und Rauheit kommt durch die **artikulatorisch-akustischen Eigenschaften des alveolaren Trills** zustande

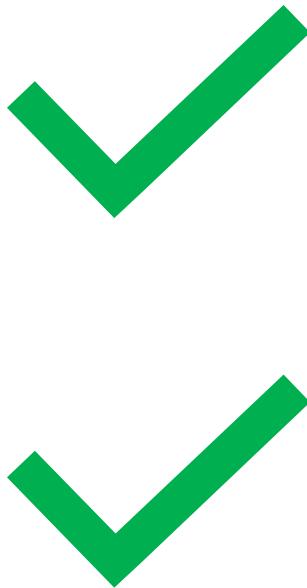


Proportion of congruent responses

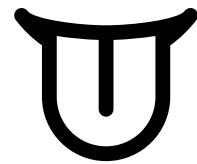
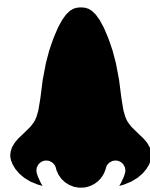
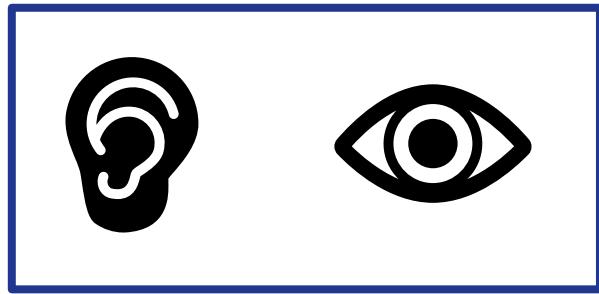


- Sehr starke Präferenz für **raue Textur bei [r] und glatte Textur bei [l]**
  - Online: 87% (Zufall: 50%)
  - Vor Ort: 97% (Zufall: 50%)
- **Kein Effekt von Phonemkontrast!**
- Im Online-Experiment:
  - Effekt schwächer in Sprachen mit [r] als Hauptrealisierung (95% vs. 85%)
  - Weil nur Sprachen ohne [r] als Hauptrealisierung vom Reihenfolgeeffekt profitieren (90% wenn [r] zuerst vs. 81% wenn [l] zuerst)

Die Wahrnehmung des phonetischen Signals ist entscheidend, aber das phonologische System kann dessen Deutlichkeit beeinflussen



Krivokapić, Tiede & Tyrone (2017); Pouw & Fuchs (2022);  
Wagner, Ćwiek & Samłowski (2019); Wagner, Malisz & Kopp (2014)



Majid et al., (2018); San Roque et al. (2014, 2018)

Akustisch-auditive Ikonizität hat das Potenzial, ein **Kommunikationssystem zu gründen**

Das **phonetische Potenzial** der Ikonizität kann durch die **Phonologie beeinflusst** werden

Multimodale Koordination ist bedeutungsrelevant

01



02

03

04



1. Graffiti



2. Feuerwerk



3. Bombe



Blitz



- 207 Konzepte, die sich in Bezug auf Sinnesbereich, Abstraktheit, Ikonizität, Valenz usw. unterscheiden
- *Wie gut könnten Sie dieses Konzept **ohne Sprache** nur mit der Stimme/nur mit Gesten/sowohl mit der Stimme als auch mit Gesten ausdrücken?*
- Bewertung auf einer Skala von 0 bis 1
- 248 deutsche und 215 niederländische Muttersprachler\*innen
- Wir würden dieses Experiment gerne sprach- und kulturübergreifend durchführen!

- Bayesian hierarchical zero-one inflated beta Modell
- 4 chains á 4000 Iterationen (insgesamt 8000 post-warmup)

```
brm(bf(expressibility_norm ~ 1 + modality +
         (1 + modality || word),
         phi ~ modality + (1 + modality || word),
         zoi ~ modality + (1 + modality || word),
         coi ~ modality + (1 + modality || word)))
```

Parameter	Estimate	SE	95% CI	p( $\beta < 0$ )	ESS
<b>Fixed Effects</b>					
Multimodal	0.35	0.03	[0.29, 0.42]		2066
Gesture	-0.11	0.03	[-0.17, -0.04]	100%	4670
Vocal	-0.81	0.04	[-0.89, -0.73]	100%	3391
<b>Random Effects (word)</b>					
sd(Multimodal)	0.74	0.04	[0.67, 0.82]		1769
sd(Gesture)	0.37	0.03	[0.31, 0.43]		2977
sd(Vocal)	0.53	0.04	[0.46, 0.60]		3821

- 84-Item-Liste mit „am besten ausdrückbaren“ Konzepten
- *Drücken Sie das Konzept so aus, dass Ihr Gegenüber es erraten kann*
- Drei verschiedene Modalitätsbedingungen: vokal, gestisch, multimodal
- Erraten durchgehend durch die\*den Wahrnehmende\*n
- Kein Feedback vs. Feedback an die\*den Produzierende\*n  
(ohne vs. mit Interaktion)

- (Analyse-)Datenbank wird OA aufgebaut  
([envisionBOX.org](https://envisionBOX.org) und [ViCom Datennetzwerk](#))
- Beispiele der Forschungsfragen:
  - „Natural statistics“: wie unterscheiden sich uni- vs. multimodale Signale? (Ćwiek & Fuchs, 2020; Macuch Silva et al., 2020)
  - Trägt die multimodale Koordination zur Prominenz bei?  
(gemessen als Erkennungsrate)
  - Wird die Erkennungsrate durch die Interaktion beeinflusst?
  - **Ökonomisieren wir in Wiederholungen?** (principle of least effort; H&H, auch AVT genannt, Lindblom, 1990)

OSF preregistration  
Kadavá, Ćwiek,  
Fuchs & Pouw (2023)



Akustisch-auditive Ikonizität hat das Potenzial, ein **Kommunikationssystem zu gründen**

Das **phonetische Potenzial** der Ikonizität kann durch die **Phonologie beeinflusst** werden

Multimodale **Koordination** ist **bedeutungsrelevant**

Sprachübergreifende Studien zeigen das Zusammenspiel von Multimodalität und Ikonizität beim Schaffen von **Bedeutung**

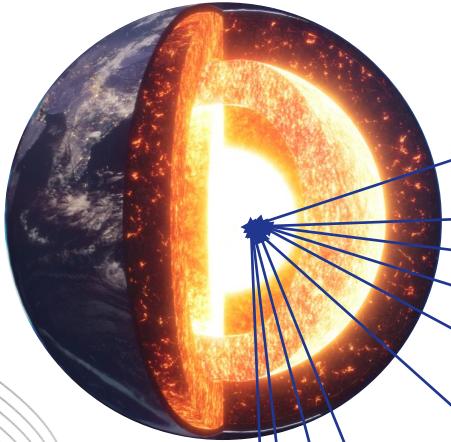
**01**

**02**

**03**

**04**





Gesten oder Gebärden? (...) (Pl: Bauer)

Multimodal feedback signals across language modalities (Pl: Bauer)

The relation between eye gaze and vocal feedback (Pl: Grice & Perniss)

The Gesture-to-Sign Trajectory: Phonological Parameters in Production and Real-Time Comprehension (Pl: Perniss & Schumacher)

Prominence in action: Referent representation in German Sign Language (DGS) (Pl: Perniss)

Loos, C., Steinbach, M., & Repp, S., (Angenommen). Polar response strategies across modalities: Evidence from German Sign Language (DGS). *Language*.

Spaniol, M., Janz, A., Wehrle, S., Vogeley, K., & Grice, M. (2023). Multimodal signalling: The interplay of oral and visual feedback in conversation. *Proceedings of the 20th ICPHS*.

Pagel, L., Sóskuthy, M., Roessig, S., & Mücke, D. (2023). A kinematic analysis of visual prosody: Head movements in habitual and loud speech. *Proceedings of the 20th ICPHS*.

Mertz, J., Pagel, L., Turco, G. & Mücke, D. (To appear). Gradiency and categoriality in the prosodic modulation of French Sign Language: A kinematic approach using Electromagnetic Articulography. *Proceedings of the International Conference on Speech Prosody 2024. (+ 19. LabPhon)*

Baills, F. & Baumann, S. (2023). The multimodal marking of information status in French as a foreign language: What can we learn about the use of prosodic and gestural cues in an interlanguage? *PaPE 2023, Nijmegen*.

# *Language is iconic and multimodal to the core*

(Perlman, 2017)

# Vielen Dank

# Referenzen

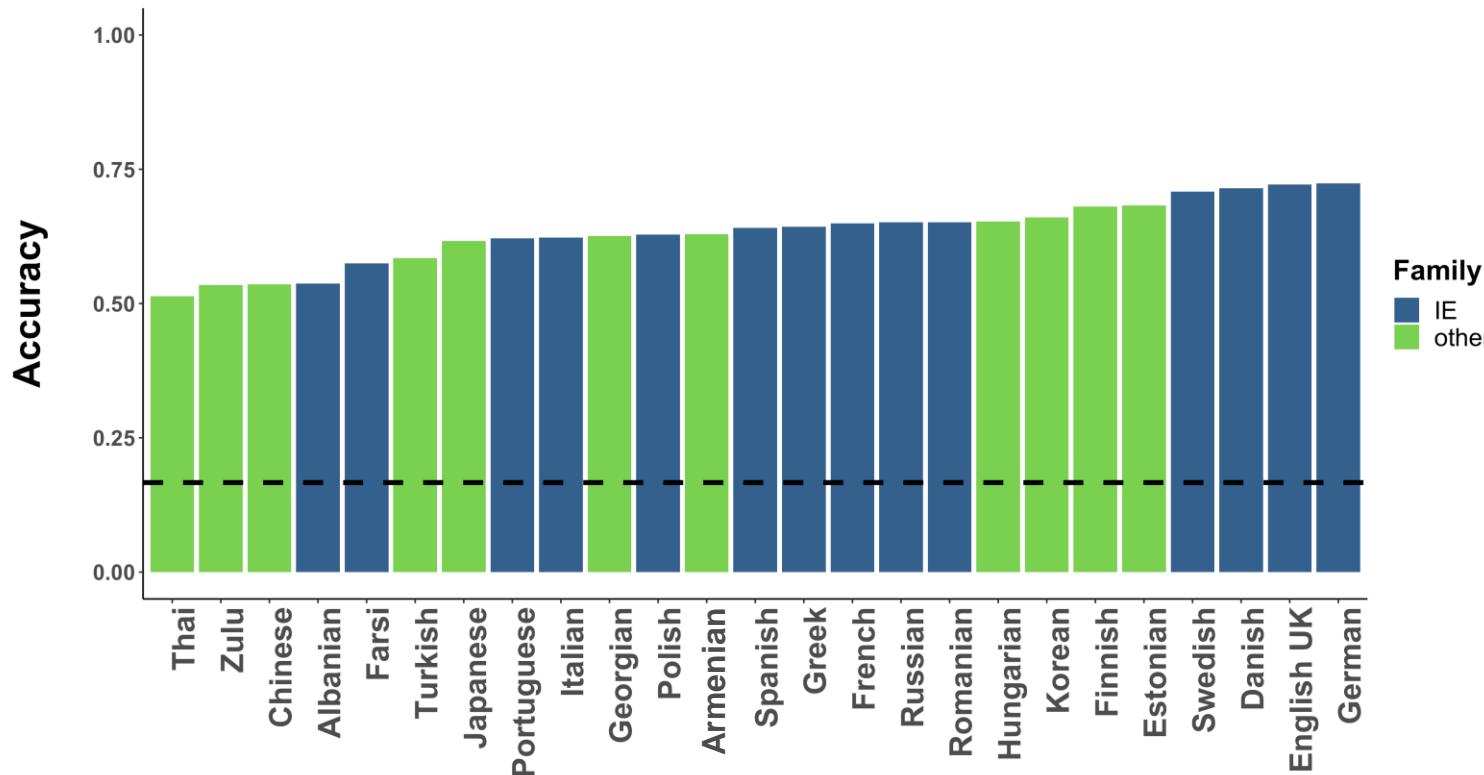
- Anselme, R., Pellegrino, F., & Dediu, D. (2022, September 5). R You Sure That Your /r/ Is Trilled? A Methodological Caveat. *Joint Conference on Language Evolution*. <https://cnrs.hal.science/hal-03900020>
- Anselme, R., Pellegrino, F., & Dediu, D. (2023). What's in the r? A review of the usage of the r symbol in the illustrations of the IPA. *Journal of the International Phonetic Association*, 1–30. <https://doi.org/10.1017/S0025100322000238>
- Ćwiek, A. (2022). Iconicity in Language and Speech. PhD Thesis, Humboldt-Universität zu Berlin. <https://doi.org/10.18452/24544>
- Ćwiek, A., & Fuchs, S. (2019). Iconic prosody is rooted in sensori-motor properties: Fundamental frequency and the vertical space. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 41, 1572–1578. <https://cogsci.mindmodeling.org/2019/papers/0282/0282.pdf>
- Ćwiek, A., Anselme, R., Kawahara, S., Perlman, M., Winter, B., Dediu, D., \* & Fuchs, S.\* (eingereicht bei JASA). The alveolar trill is perceived rough by speakers of different languages. (\* gleichwertiger Beitrag)
- Ćwiek, A., Fuchs, S., Draxler, C., Asu, E. L., Dediu, D., Hiovain, K., Kawahara, S., Koutalidis, S., Krifka, M., Lippus, P., Lupyan, G., Oh, G. E., Paul, J., Petrone, C., Ridouane, R., Reiter, S., Schümchen, N., Szalontai, Á., Ünal-Logacev, Ö., ... Perlman, M. (2021). Novel vocalizations are understood across cultures. *Scientific Reports*, 11(1), 10108. <https://doi.org/10.1038/s41598-021-89445-4>
- Ćwiek, A., & Fuchs, S. (2020). Hand-Mouth Coordination in a Pointing Task Requiring Manual Precision. *Proceedings of the 12th ISSP*. 12th International Seminar on Speech Production, New Haven, CT: Haskins Laboratories.
- de Saussure, F. (1916). *Cours de linguistique générale*.
- Dingemanse, M. (2012). Advances in the Cross-Linguistic Study of Ideophones. *Language and Linguistics Compass*, 6(10), 654–672. <https://doi.org/10.1002/lnc.3.361>
- Freywald, U. (2016). Total reduplication as a productive process in German. *The Why and How of Total Reduplication: Current Issues and New Perspectives*, 39(4), 905–945. <https://doi.org/10.1075/sl.39.4.06fre>
- Fuchs, S., & Ćwiek, A. (2022). Sounds Full of Meaning and the Evolution of Language. *Acoustics Today*, 18(2), 43–51. <https://doi.org/10.1121/AT.2022.18.2.43>
- Gardner, R. A., & Gardner, B. T. (1969). Teaching Sign Language to a Chimpanzee. *Science*, 165(3894), 664–672.
- Hewes, G. W. (1973). Primate Communication and the Gestural Origin of Language. *Current Anthropology*, 14(1/2), 5–24.

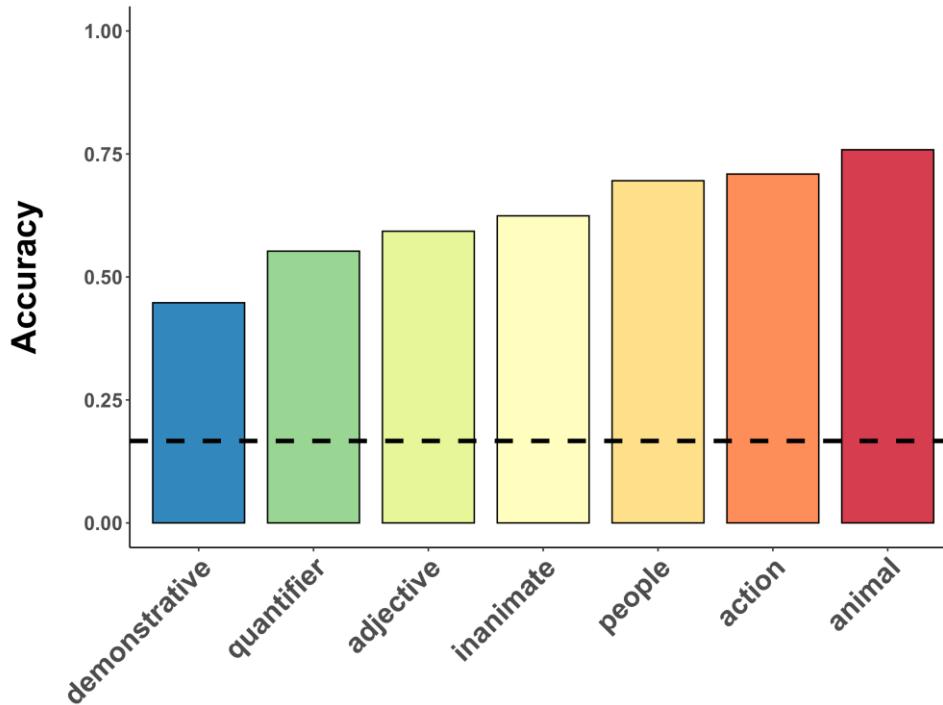
# Referenzen

- Himmelmann, N. P. (2005). The Austronesian languages of Asia and Madagascar: Typological characteristics. *The Austronesian languages of Asia and Madagascar*, 110, 110-181.
- Hockett, C. F. (1960). The origin of speech. *Scientific American*, 203(3), 88–96. <https://doi.org/10.1038/scientificamerican0960-88>
- Hockett, C. F. (1978). In search of Jove's brow. *American Speech*, 53(4), 243–313. <https://doi.org/10.2307/455140>
- Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651), 20130298. <https://doi.org/10.1098/rstb.2013.0298>
- Kadavá, Š., Ćwiek, A., Fuchs, S., & Pouw, W. (2023). Pre-registration: The Coordination of Dynamic Multimodal Signals in Novel Communication. OSF. <https://osf.io/https://osf.io/bdpvy>
- Kadavá, Š., Ćwiek, A., Fuchs, S., & Pouw, W. (2024). What do we mean when we say gestures are more expressive than vocalizations? An experimental and simulation study. *Proceedings of CogSci 2024*.
- Kendon, A. (2017). Reflections on the “gesture-first” hypothesis of language origins. *Psychonomic Bulletin & Review*, 24(1), 163–170. <https://doi.org/10.3758/s13423-016-1117-3>
- Kilpatrick, A., & Ćwiek, A. (2024). Using artificial intelligence to explore sound symbolic expressions of gender in American English. *PeerJ Computer Science*, 10, e1811. <https://doi.org/10.7717/peerj-cs.1811>
- Krivokapić, J., Tiede, M. K., & Tyrone, M. E. (2017). A Kinematic Study of Prosodic Structure in Articulatory and Manual Gestures: Results from a Novel Method of Data Collection. *Laboratory Phonology*, 8(1). <https://doi.org/10.5334/labphon.75>
- Lindblom, B. (1990). Explaining Phonetic Variation: A Sketch of the H&H Theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modelling* (pp. 403–439). Springer Netherlands. [https://doi.org/10.1007/978-94-009-2037-8\\_16](https://doi.org/10.1007/978-94-009-2037-8_16)
- Macuch Silva, V., Holler, J., Ozuyurek, A., & Roberts, S. G. (2020). Multimodality and the origin of a novel communication system in face-to-face interaction. *Royal Society Open Science*, 7(1), 182056. <https://doi.org/10.1098/rsos.182056>
- Majid, A., Roberts, S. G., Cilissen, L., Emmorey, K., Nicodemus, B., O’Grady, L., Woll, B., LeLan, B., Sousa, H. de, Cansler, B. L., Shayan, S., Vos, C. de, Senft, G., Enfield, N. J., Razak, R. A., Fedden, S., Tufvesson, S., Dingemanse, M., Ozturk, O., ... Levinson, S. C. (2018). Differential coding of perception in the world’s languages. *Proceedings of the National Academy of Sciences*, 115(45), 11369–11376. <https://doi.org/10.1073/pnas.1720419115>

# Referenzen

- Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound symbolism* (pp. 325–347). Cambridge University Press. <https://doi.org/10.1017/cbo9780511751806.022>
- Perlman, M. (2017). Debunking two myths against vocal origins of language. *Interaction Studies*, 18(3), 376–401. <https://doi.org/10.1075/is.18.3.05per>
- Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: From a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651), 1–14. <https://doi.org/10.1098/rstb.2013.0300>
- Perry, L. K., Perlman, M., & Lupyan, G. (2015). Iconicity in English and Spanish and its relation to lexical category and age of acquisition. *PLOS ONE*, 10(9), e0137147. <https://doi.org/10.1371/journal.pone.0137147>
- Pouw, W., & Fuchs, S. (2022). Origins of vocal-entangled gesture. *Neuroscience & Biobehavioral Reviews*, 141, 104836. <https://doi.org/10.1016/j.neubiorev.2022.104836>
- San Roque, L., Kendrick, K. H., Norcliffe, E., Brown, P., Defina, R., Dingemanse, M., Dirksmeyer, T., Enfield, N. J., Floyd, S., Hammond, J., Rossi, G., Tufvesson, S., van, P. S., & Majid, A. (2014). Vision verbs dominate in conversation across cultures, but the ranking of non-visual verbs varies. *Cognitive Linguistics*, 26(1), 31–60. <https://doi.org/10.1515/cog-2014-0089>
- San Roque, L., Kendrick, K. H., Norcliffe, E., & Majid, A. (2018). Universal meaning extensions of perception verbs are grounded in interaction. *Cognitive Linguistics*, 29(3), 371–406. <https://doi.org/10.1515/cog-2017-0034>
- Wagner, P., Ćwiek, A., & Samłowski, B. (2019). Exploiting the speech-gesture link to capture fine-grained prosodic prominence impressions and listening strategies. *Journal of Phonetics*, 76, 100911. <https://doi.org/10.1016/j.wocn.2019.07.001>
- Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication*, 57, 209–232. <https://doi.org/10.1016/j.specom.2013.09.008>
- Winter, B., Sóskuthy, M., Perlman, M., & Dingemanse, M. (2022). Trilled /r/ is associated with roughness, linking sound and touch across spoken languages. *Scientific Reports*, 12(1), 1035. <https://doi.org/10.1038/s41598-021-04311-7>
- Winter, B., Woodin, G., & Perlman, M. (2023). Defining iconicity for the cognitive sciences. *Preprint; for Oxford Handbook of Iconicity in Language*. <https://doi.org/10.31219/osf.io/5e3rc>



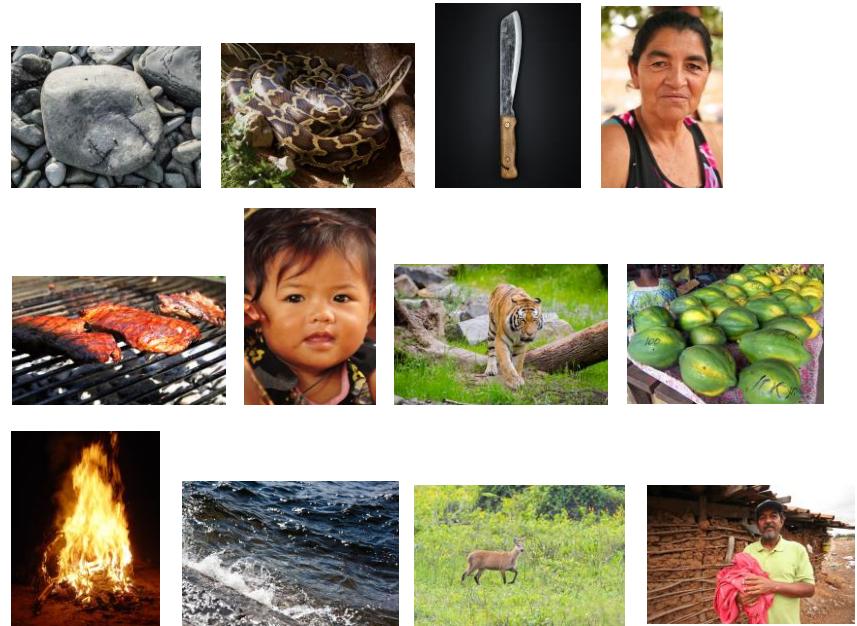


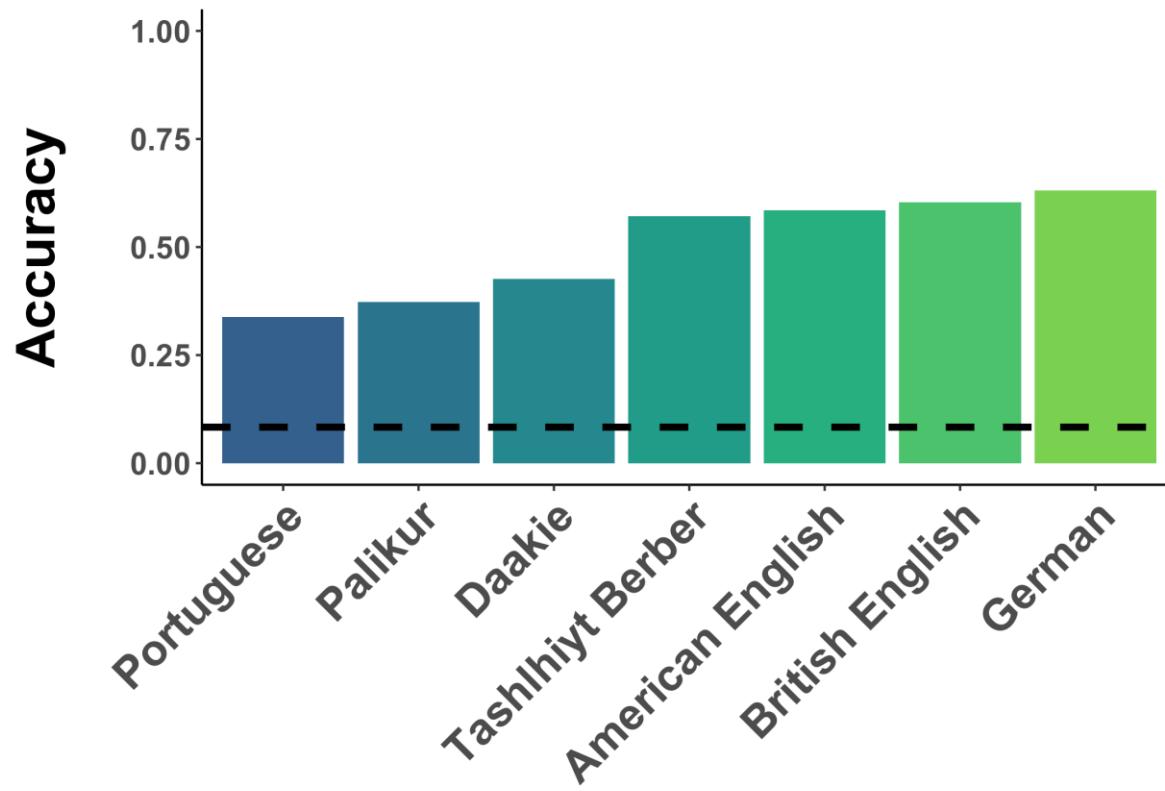
Verben sind ikonischer also  
Nomen und Adjektive:

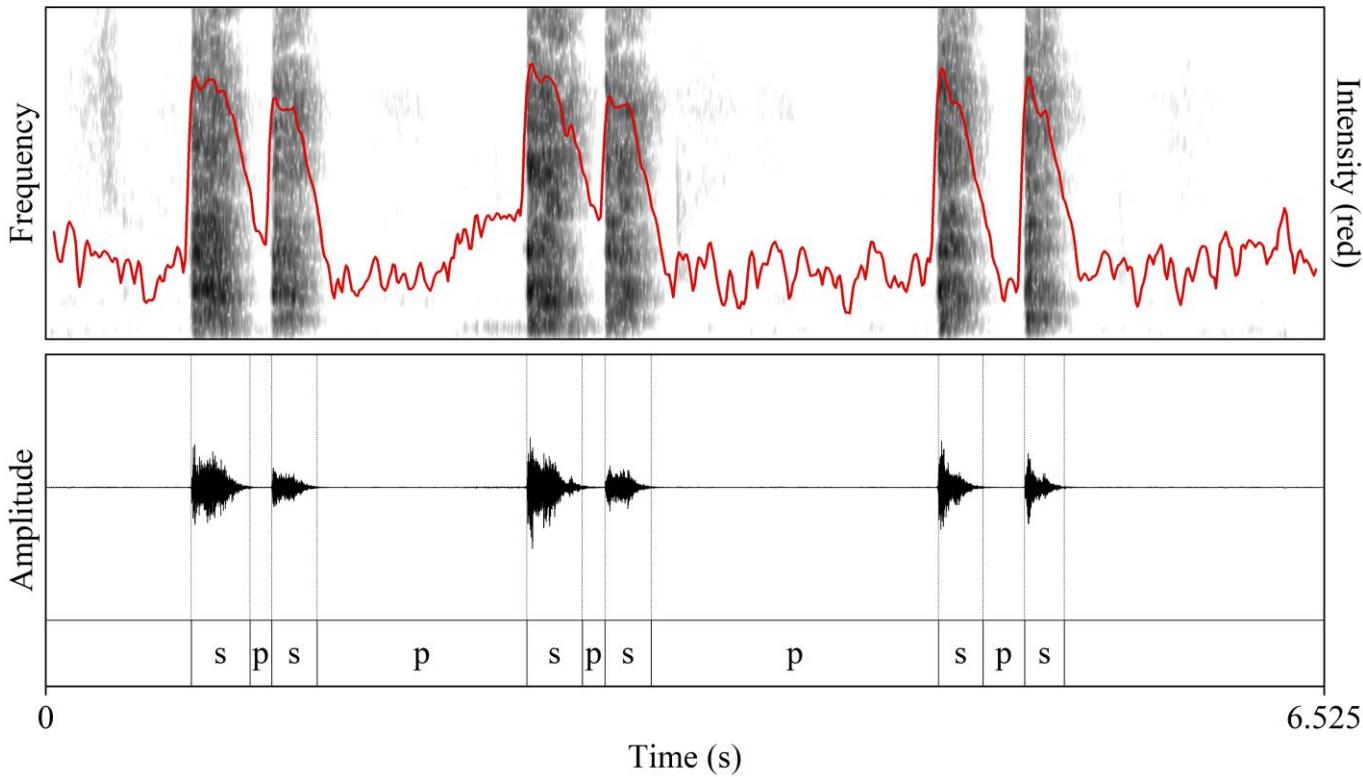
- Perry et al. (2015)
- Winter et al. (2017)
- Perry et al. (2018)
- Perlman et al. (2018)

# Vor-Ort-Experiment

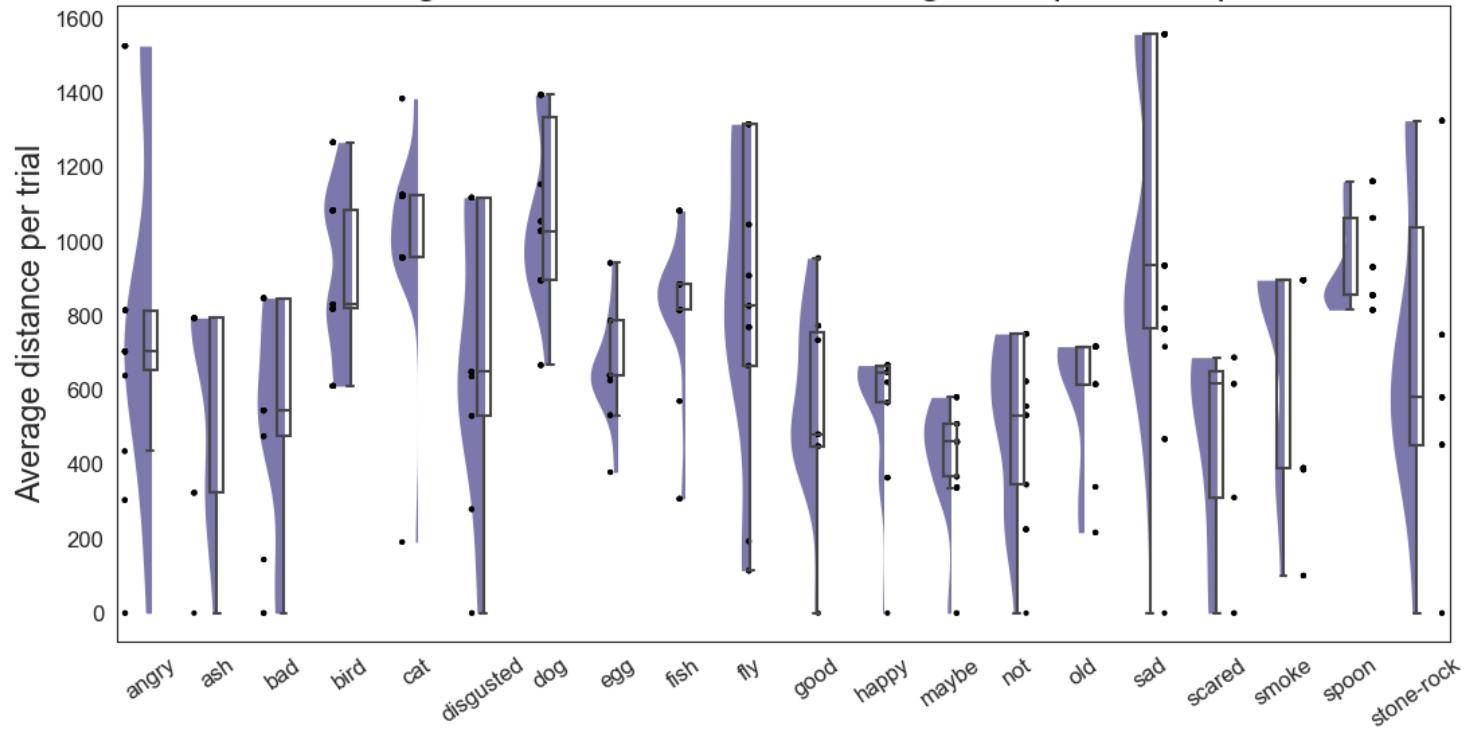
- Entwickelt für abgelegene, durchaus nicht alphabetisierte Bevölkerungsgruppen
- Nur die 3 wichtigsten Vokalisierungen für jedes Substantiv ( $3 \times 12 = 36$  Vokalisierungen)
- Die Teilnehmer zeigten auf eines von 12 Bildern, die jedes Substantiv darstellten

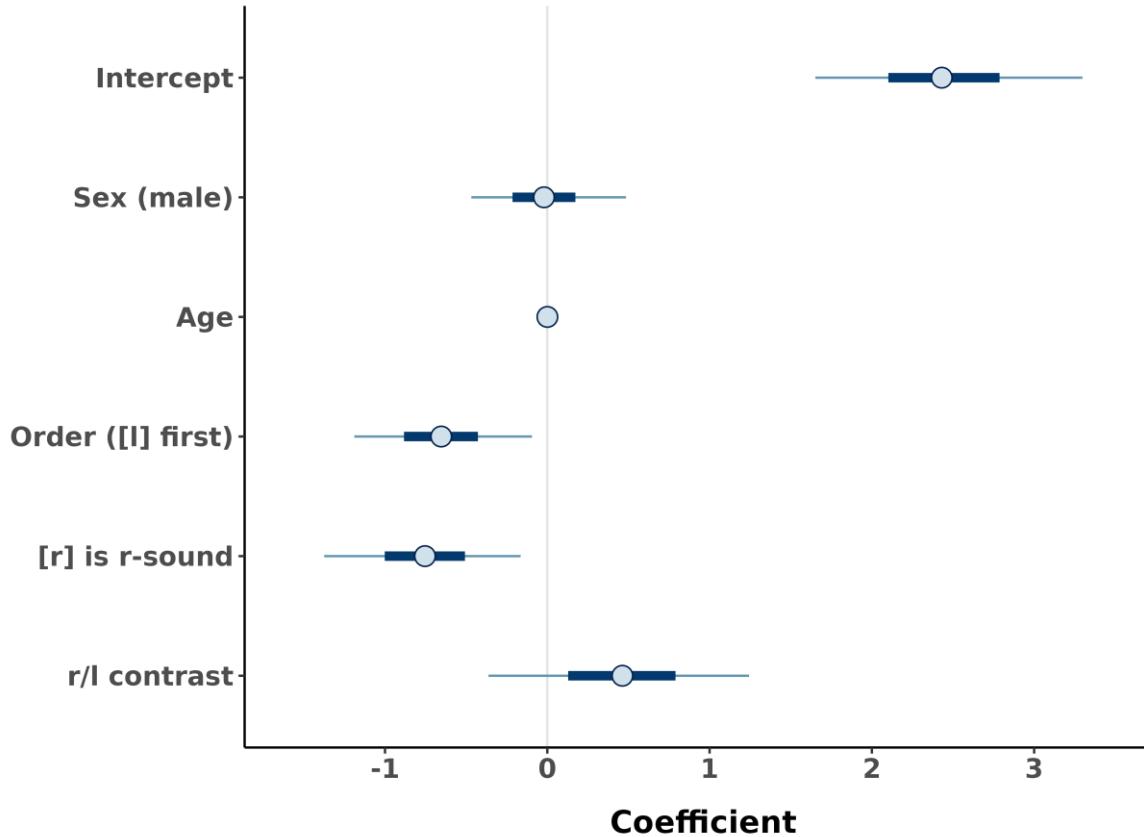


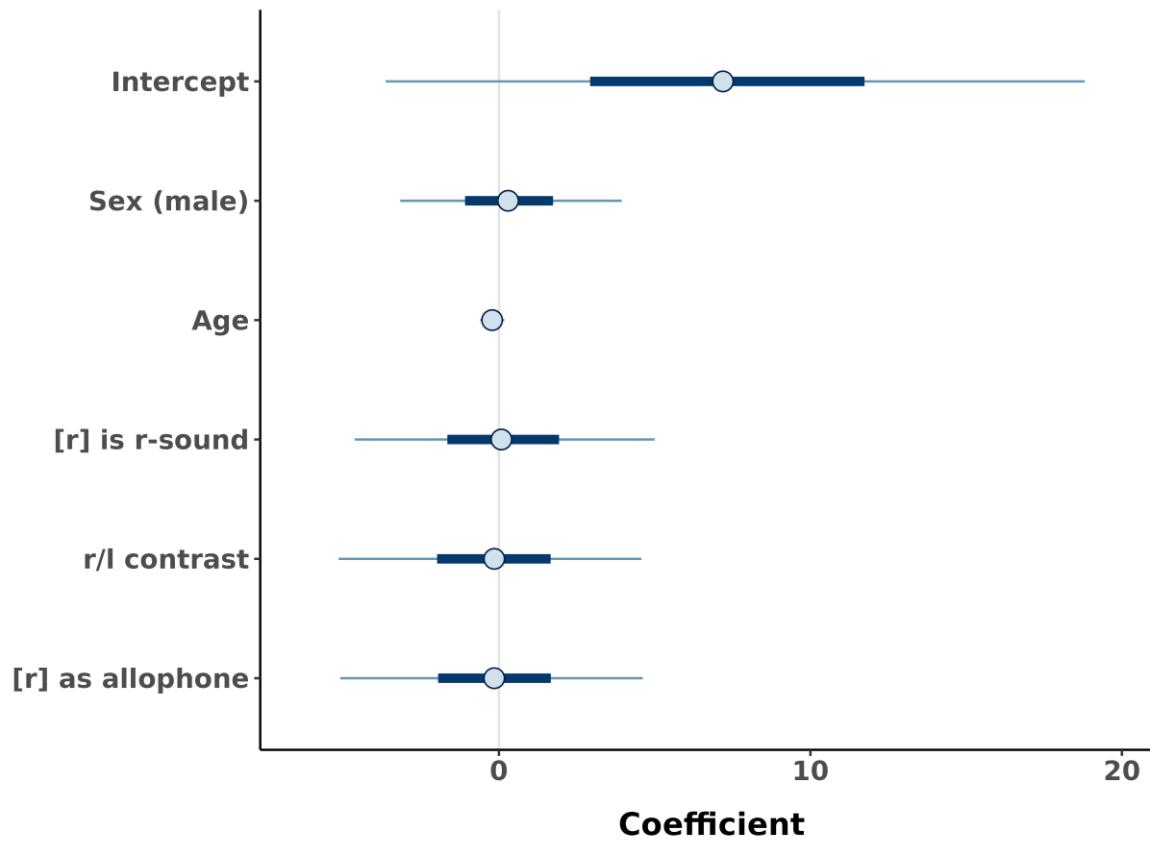


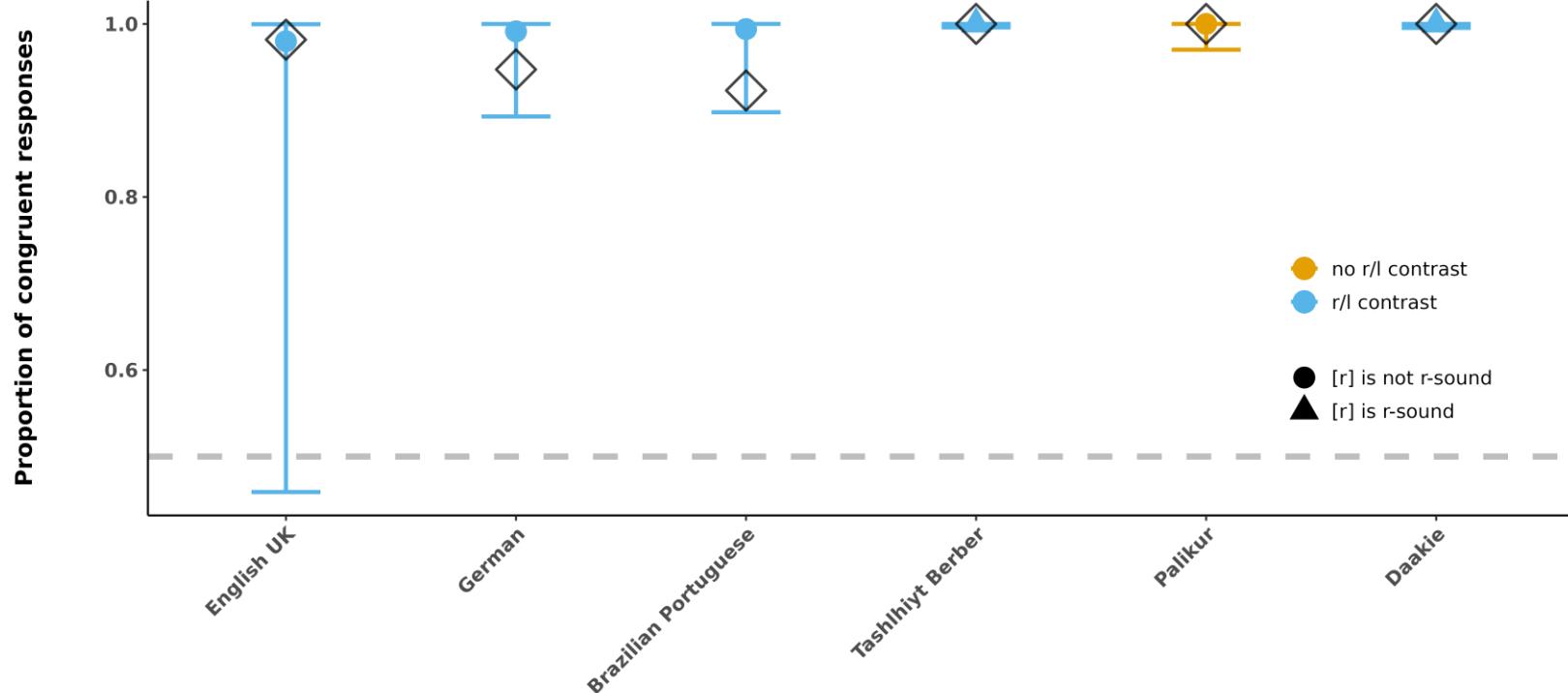


### Average distance between sound segments per concept









SemanticCat	SemanticSubcat	PoS	Info	German	English
environment	environment	adjective	adjective; opposite of wet	trocken	dry
sensory	auditory	noun	noun; singular; sound of bee	Summen	buzz
			adjective; opposite of silent; disturbing noise		
sensory	auditory	adjective	disturbing noise	geräuschvoll	noisy
sensory	auditory	adjective	adjective; opposite of noisy	leise	silent
sensory	auditory	verb	verb (trans.)	flüstern	to whisper
sensory	gustatory	noun	noun; singular	Honig	honey
sensory	gustatory	adjective	adjective	sauer	sour
sensory	gustatory	verb	verb (trans.)	kauen	to chew
sensory	gustatory	verb	verb (trans.)	lecken	to lick
sensory	gustatory	verb	verb (intrans.)	erbrechen	to vomit
sensory	haptic	adjective	adjective	eng	tight
sensory	haptic	verb	verb (trans.)	fühlen	to feel
sensory	olfactory	verb	verb (intrans.)	atmen	to breathe
			adjective; referring to a pleasant smell		
sensory	olfactory	adjective	smell	duftend	fragrant
sensory	visual	adjective	adjective; opposite of dark	hell	bright
sensory	visual	adjective	adjective; one of the color terms	gelb	yellow
abstract	logical	pronoun	pronoun; related to things	alles	all
abstract	logical	adjective	adjective; opposite of same	anders	different
abstract	logical	noun	noun; singular; opposite of past	Zukunft	future
abstract	logical	adverb	adverb	vielleicht	maybe
abstract	logical	noun	noun; singular; opposite of future	Vergangenheit	past
abstract	logical	adjective	adjective; opposite of different	gleich	same
			pronoun; describing an amount of things (uncountable/countable)		
abstract	logical	pronoun	einige		some

