

Bias for Language Simplicity in Composite Populations?

Peeter Tinitis

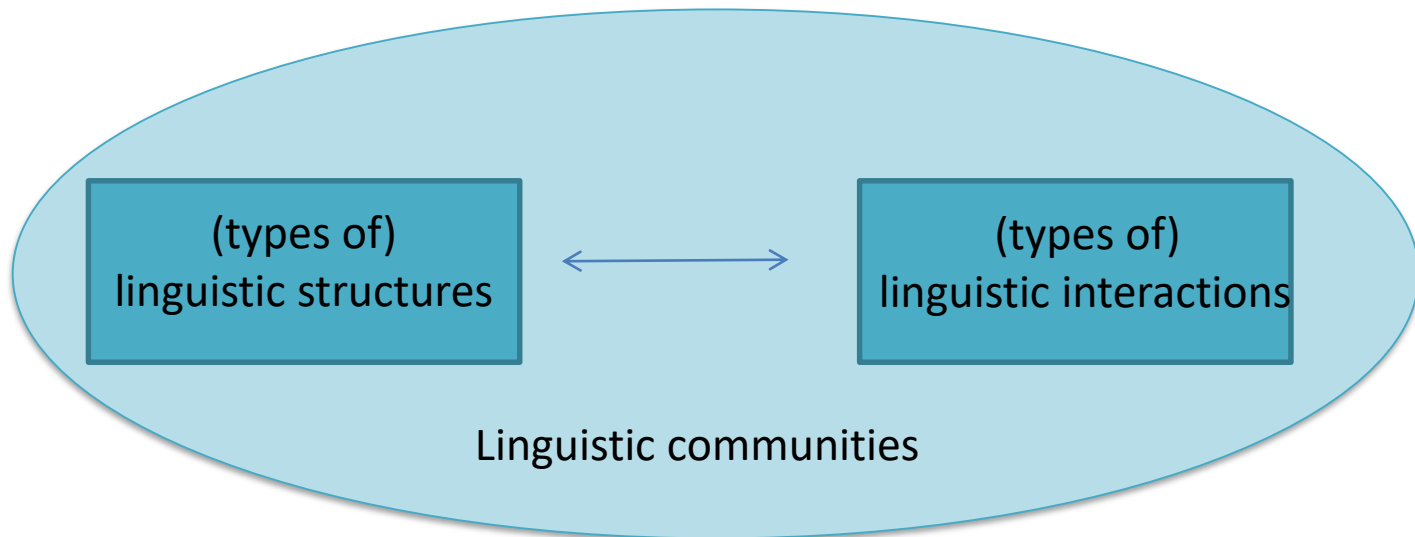
peeter.tinitis@gmail.com

Outline

- Relevance for language study
- Background for the experimental paradigm
- Experiment
 - Setup
 - Results
 - Discussion

The Issue

Hypothesis in the literature:



An example typology

Esoteric cultures

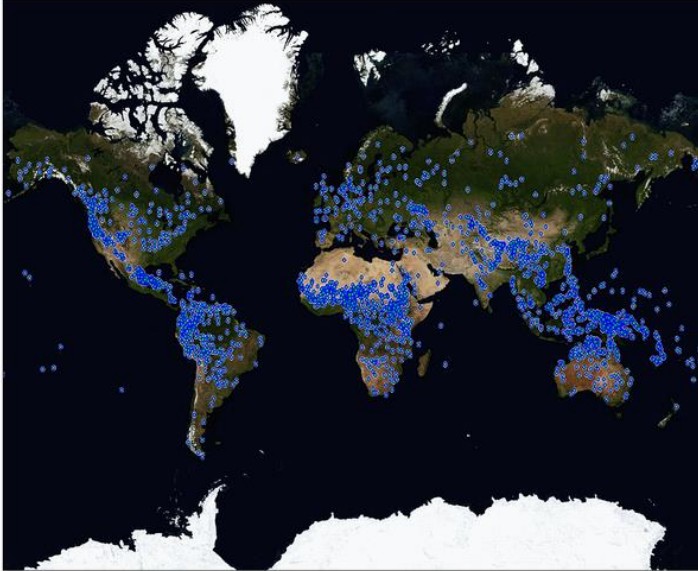
- Small, compact community
- Much shared background among the members of the community, which is also stable across generations
- Little communication with the outside world
- Closed community with little immigration
- Structural forms may also be considered culturally significant or sacred

Exoteric cultures

- Large, sparsely connected
- Little shared background among the members of the community
- Intensive communication with other communities
- Lots of immigration and L2-speakers
- Language mostly used for practical purposes

(Thurston, 1987; Kusters, 2003)

Broad correlations



simpler inflectional morphology
(number of cases, synthesis, agreement)
or lexical strategies
(evidentiality, negation, aspect, and possession)

~

exotericity(population, area, # of neighbours)

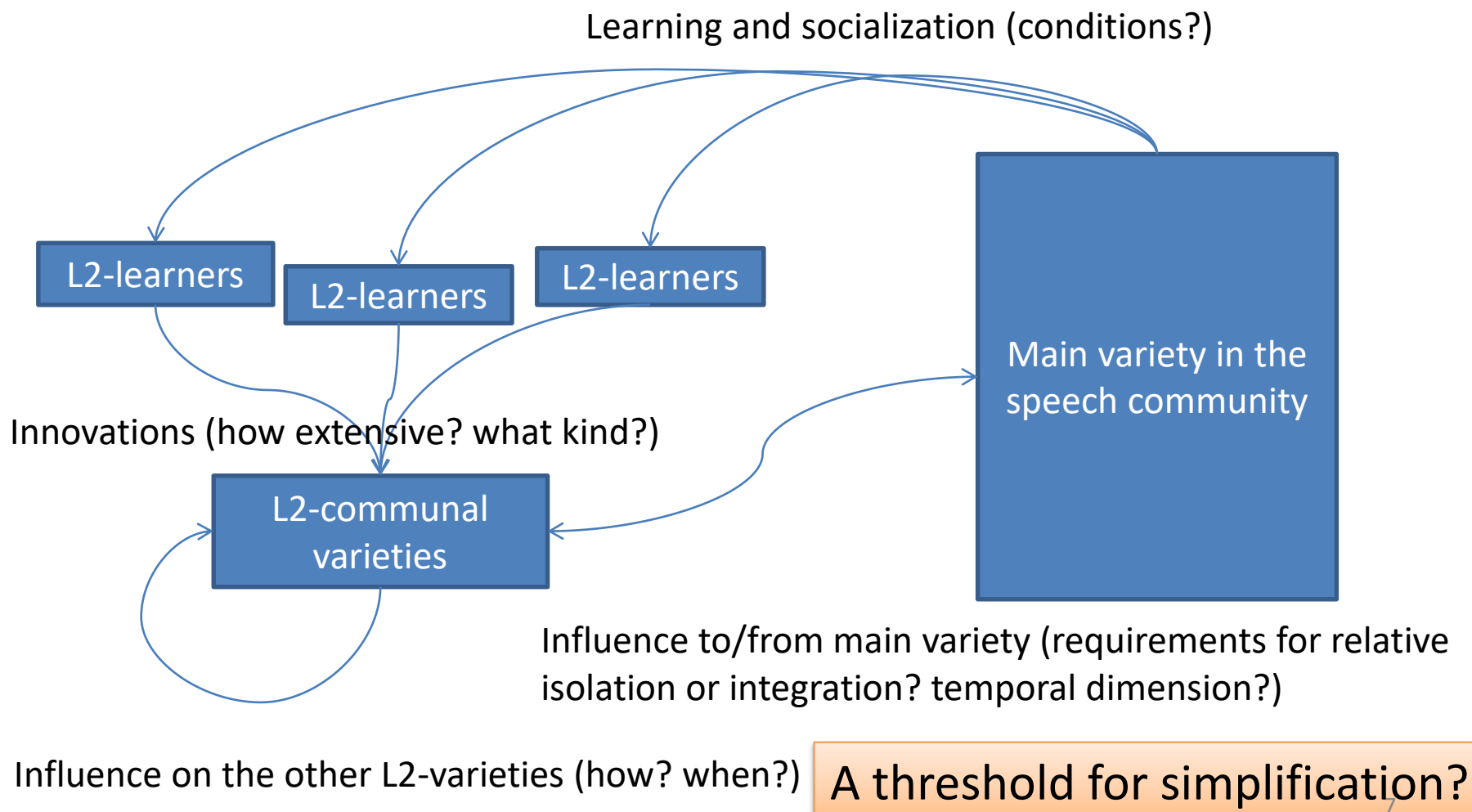
2236 languages tested for correlations by Lupyan & Dale, 2010
based on WALS and demographic sources

More esoteric languages (pop., area, # of neighbors) were significantly more likely to have either a simpler inflectional morphology or to invoke lexical strategies instead of morphological ones to encode evidentiality, negation, aspect, possession.

Dynamics of Language Simplification

- **What are the precise mechanisms by which the relative openness of the society may influence the evolution of linguistic structure?**
- McWhorter (2007: 101) – communities of L2-learners sediment on simplified varieties which then spread to the rest of the society due to sheer numbers/prestige
- Dahl (2004: 128, 281) – L2-learners in suboptimal conditions can overgeneralize features which are difficult to acquire, large groups of L2-learners may influence the structure without any language shift taking place
- Kusters (2003: 367) – as language rapidly expands to new domains (e.g. geographical), where L2-speakers use it for communication purposes in a setting isolated and long-lasting enough for the new variety to establish itself in its own right, a simplified variety can emerge

Simple model of simplification

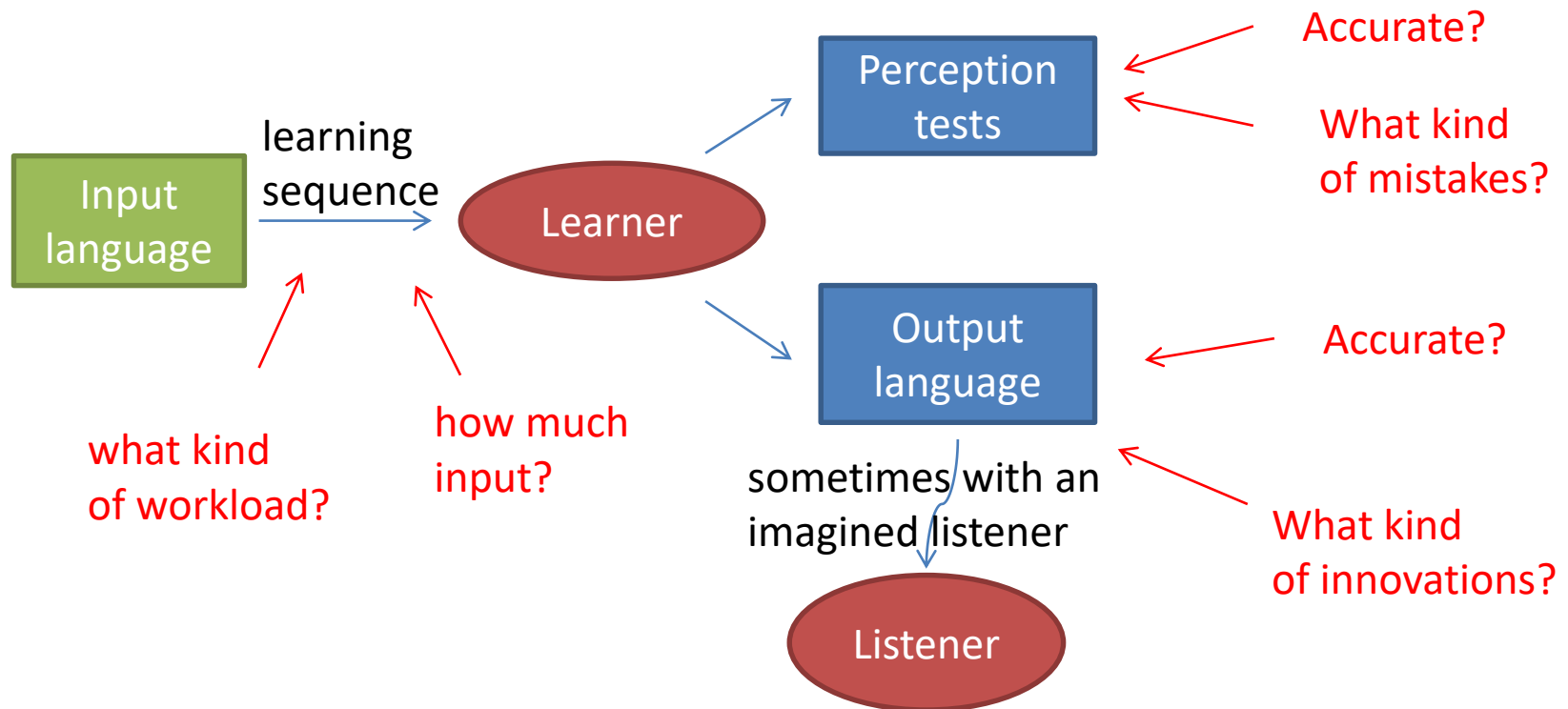


Experiments with artificial languages

- An experimental proxy for features of interest in natural languages
 - learnable in limited time
 - easily manipulable to represent the structural properties in the focus of research
 - allows good control over individual language experience and dialectal background
- Open issues
 - miniature languages -> full languages ?
 - experimental contexts -> natural contexts ?
 - possible native language interference?

A typical ALL study

- Learning -> Production/Perception



Some experimental results

Learning input -> production & perception output

- Learners reproduced more accurately more systematic miniature languages and to a small extent innovated to increase systematicity (A. Smith et al., 2010).
- Given an unsystematic language, a linear chain of learners, where a learner learns from the output of the previous learner, introduced strong systematicity to the system as a cumulative result of language transmission (Kirby et al., 2008).
- A chain of adult learners led to a regularization of arbitrary unconstrained variation in plural markers (K. Smith & Wonnacott, 2010).
- Learners reproduced more accurately typologically most common S-O word-orders than alternatives (Greenberg universal 1, Tily et al., 2011)
- Learners regularized variation to match word order harmony (N-Adj & N-Num / Adj-N & Num-N) as is also typologically common (Greenberg universal 18, Culbertson & Smolensky, 2009)

Next step

- Questions less addressed
 - what about form diffusion?
 - what about variety formation?
 - what about contact situations?
- Answers would also impact the amount of innovations required through learning or the required presence of L2-learners for widespread language simplification.

Components of the paradigm

Traditional ALL components

- Learning sequence involving observation, perception and production
- Language is limited to small vocabulary
- Form-meaning mapping through image-word connections
- Matching game scenario for perception and production
- Form-meaning systematicity is measured against randomized form-meaning allocations

Small innovations

- Learning sequence extended to achieve competence
- Matching game placed in an interactive situation of live communication
- The use of spoken language on interactions (also Kalnins, 2010)
- Closed groups of interactions - allows for group effects in variety formation
- Initial baseline is set as historically plausible semi-systematic language

The study

- „Language Simplification in Composite Populations“
(sup. S. Kirby, H. Cornish)



- Prelearned miniature languages
- Coordination game for communities containing two similar but non-identical language varieties
- Tested for a bias for systematicity

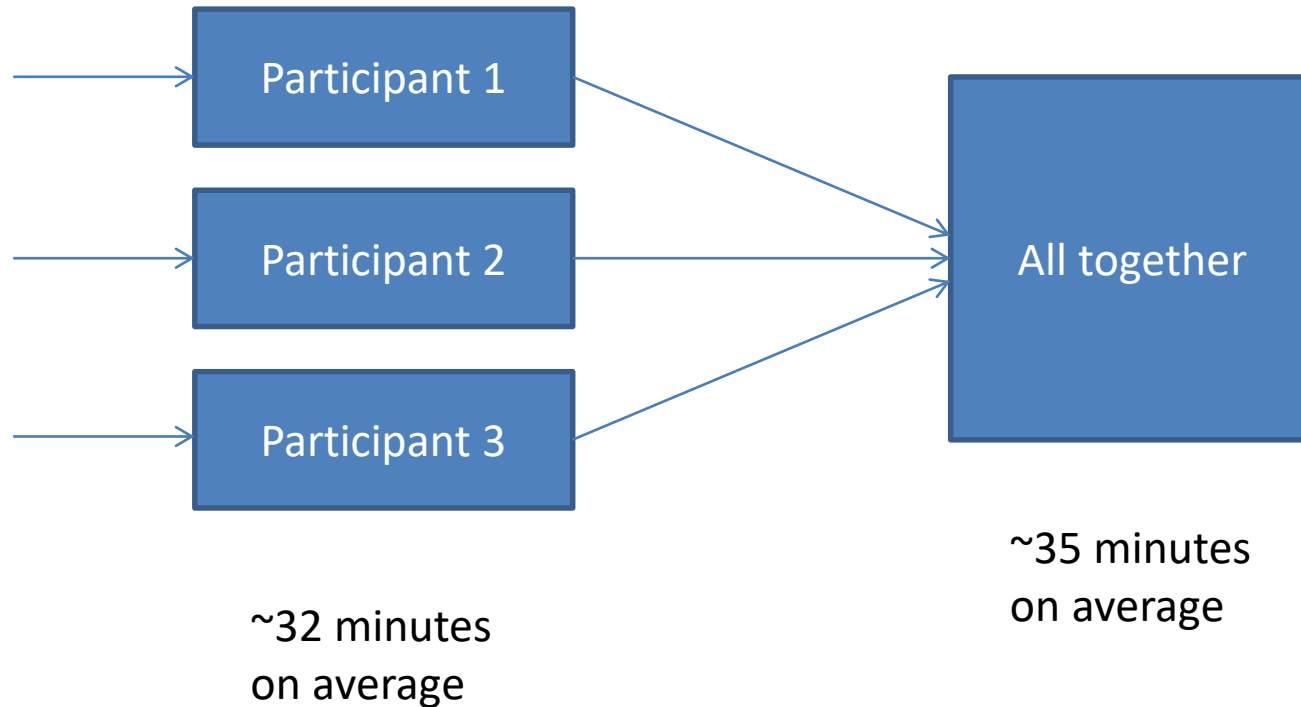
Participants

- 18 participants (6 groups of 3 people)
 - Native English speakers
 - No impairments
 - No linguists
 - Unfamiliar within the group

Experiment in 2 parts

1. Solitary learning sequence

2. Communal interaction game



Languages

- 9 items per language (3x3 meaning space)
- 2 languages in each group
 - Minority
 - Majority
- 2 types of varieties
 - Simple
 - Complex
- Minority and majority differed in 2 items

Two conditions

Simple Condition

Subject 1 – Simple language

Subject 2 – Complex language

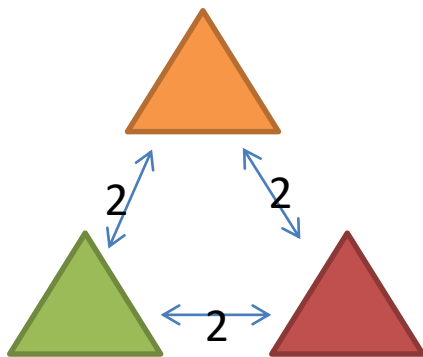
Subject 3 – Complex language

Complex Condition

Subject 1 – Complex language

Subject 2 – Complex language

Subject 3 – Complex language

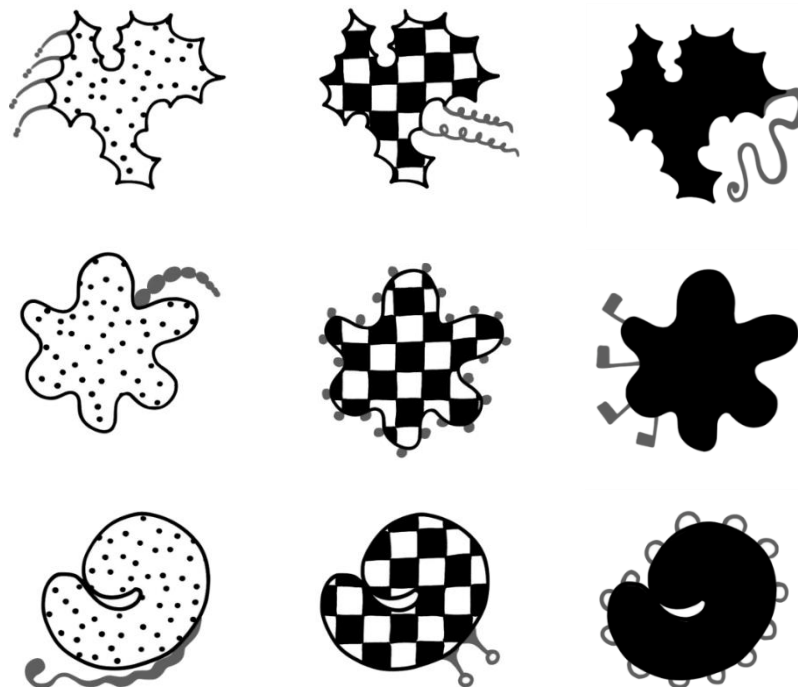


↑
Systematicity/
Simplicity

*Languages were created in sets of three, where each differed from the others by **two** maximally different items. Later they were compared pairwise.

Meanings

- From a study by Tamariz et al. (2012)
- 3 x 3 interrelated
- Systematically related, yet idiosyncratic

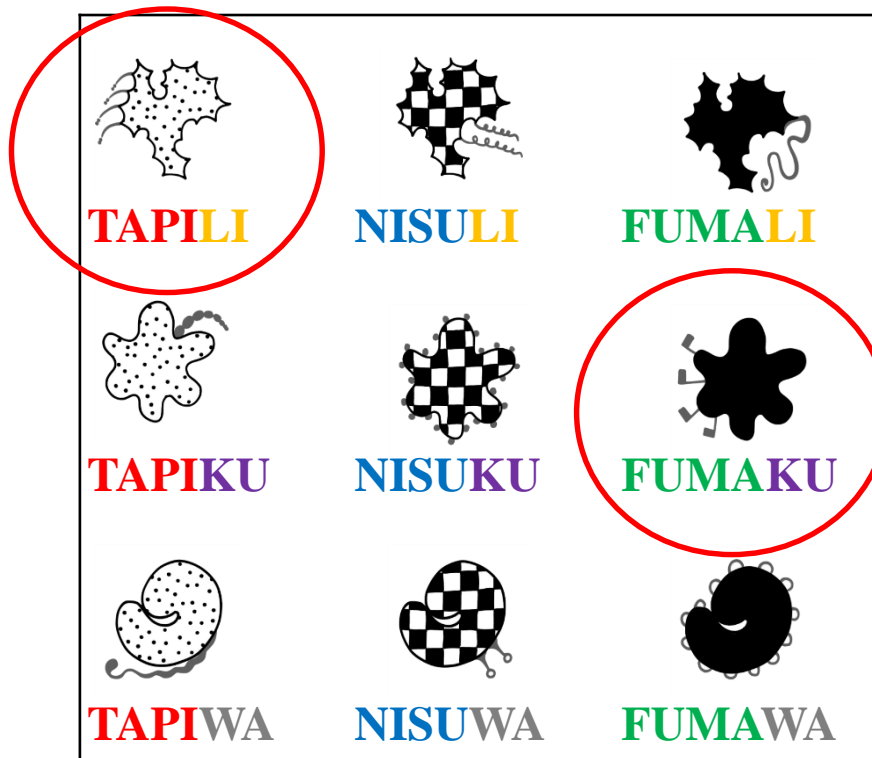


Form

- 3 x 3 (semi)systematic vocabulary of 9 words
- Limited phoneme space in the training items
- 9 consonants, 3 vowels (/k/, /p/, /t/, /s/, /m/, /n/, /l/, /f/, /w/, /a/, /i/, /u/)
- CV-CV-CV structure
- Tested against chance similarities

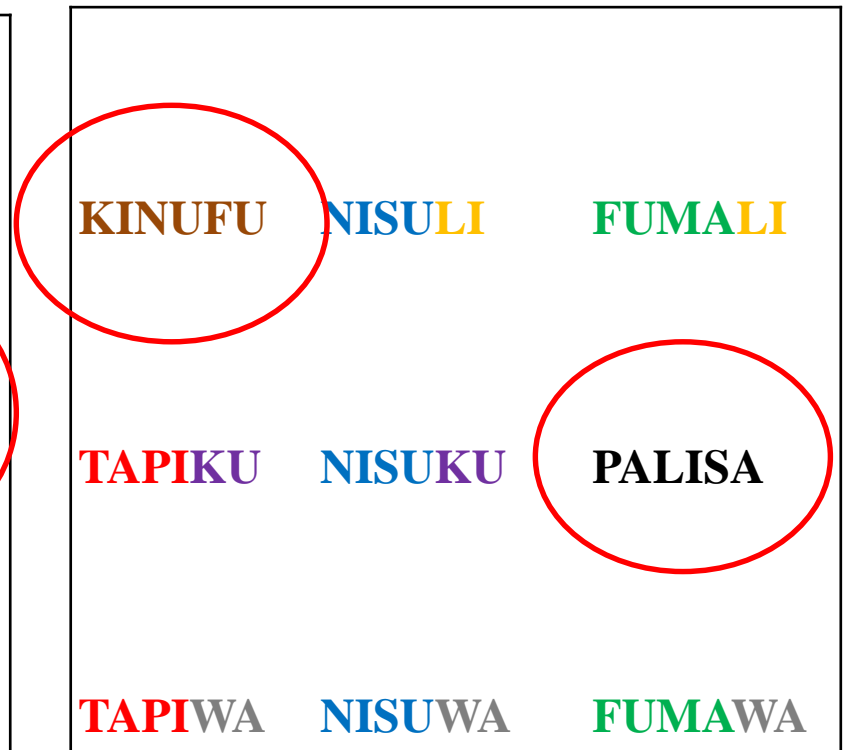
Simple Condition

Minority language



Simplicity score: 4.66

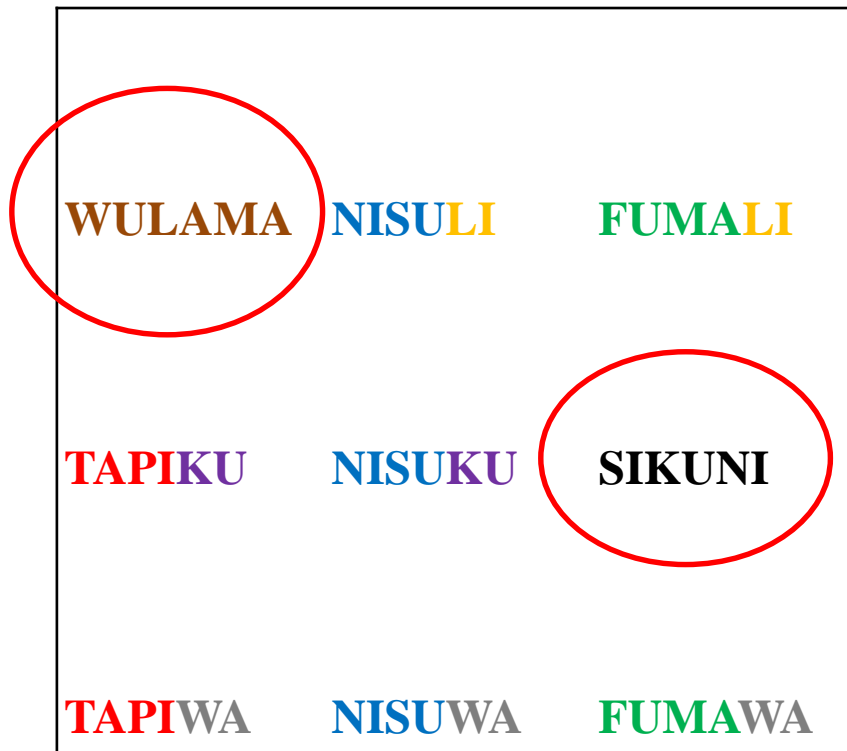
Majority language



Simplicity score: 2.89

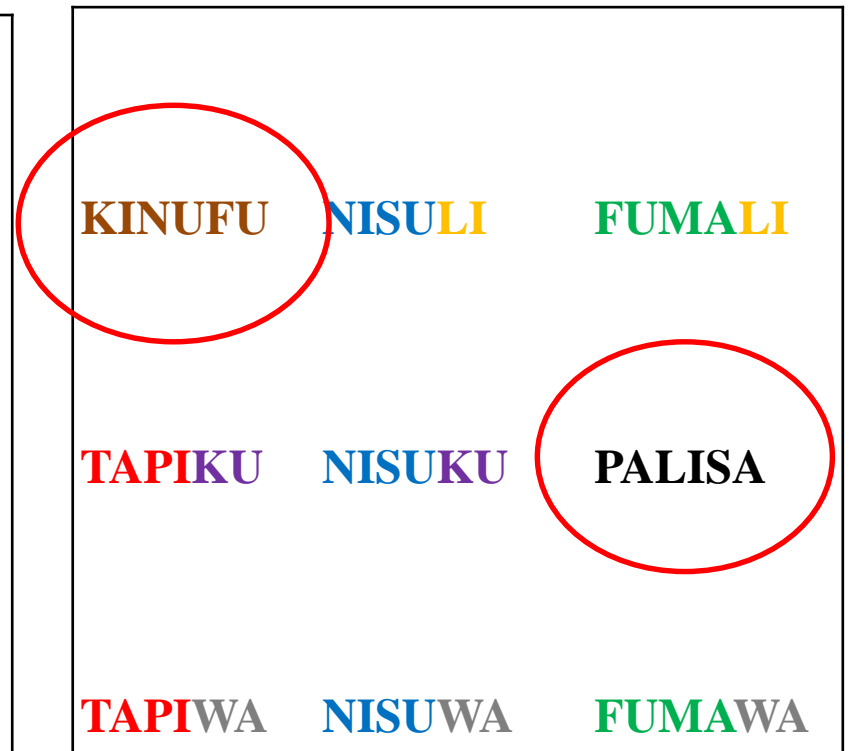
Complex Condition

Minority language



Simplicity score: 2.87

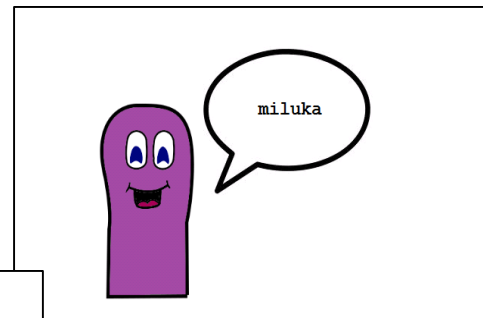
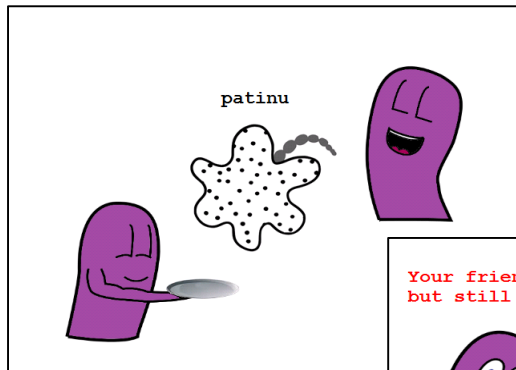
Majority language



Simplicity score: 2.89

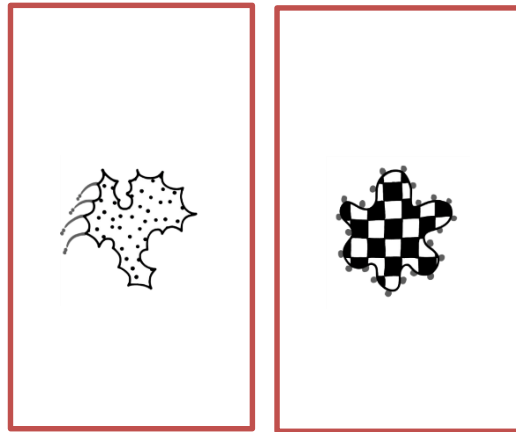
Learning

- 8 rounds = equal # of trials for all participants
 - (192 x production, 192 x perception, 192 x observation interweaved)
 - (2 obs -> 1 perc -> 2 prod -> 1 perc ... repeats)
- Expected to be competent by the end

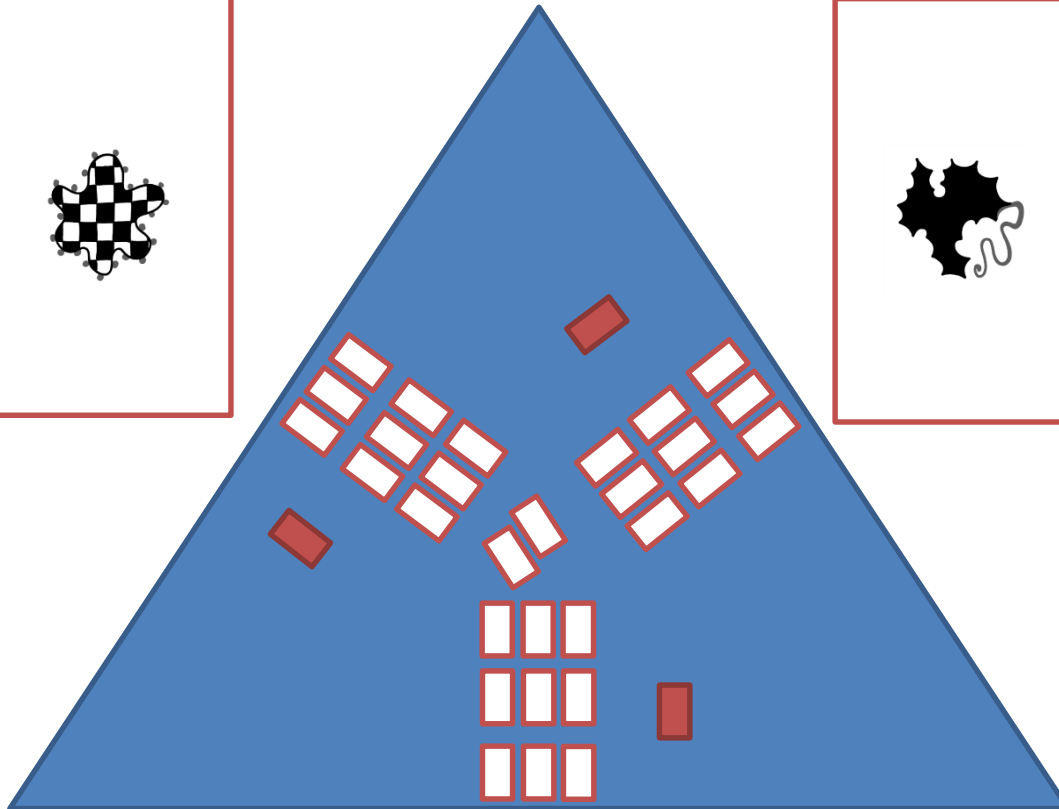
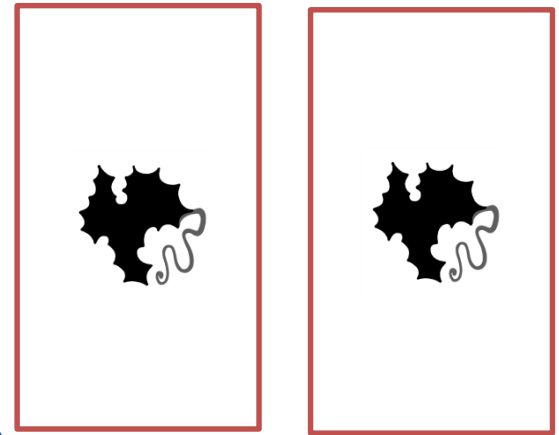


The Game

Mismatch

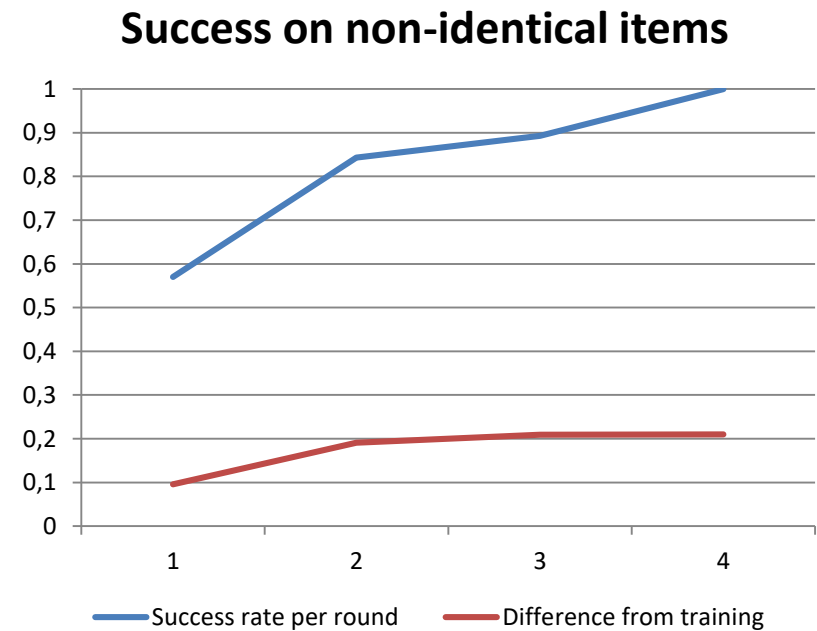
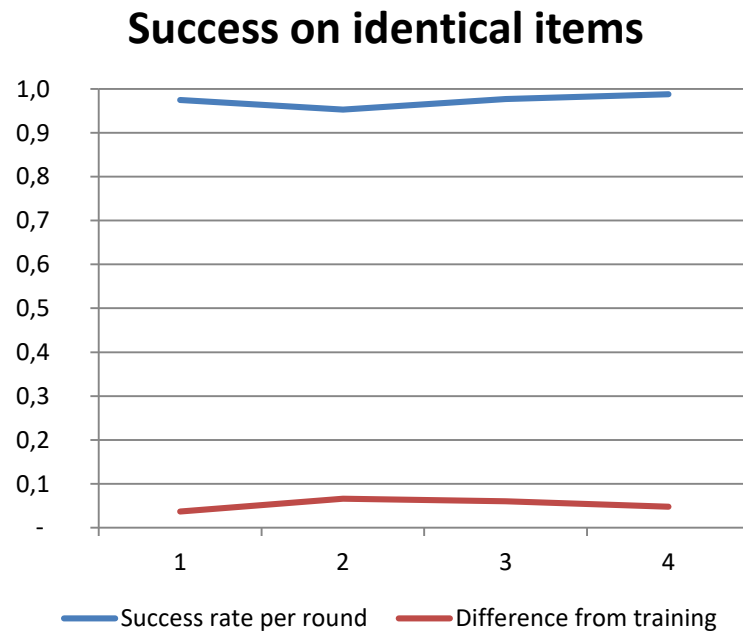


Match



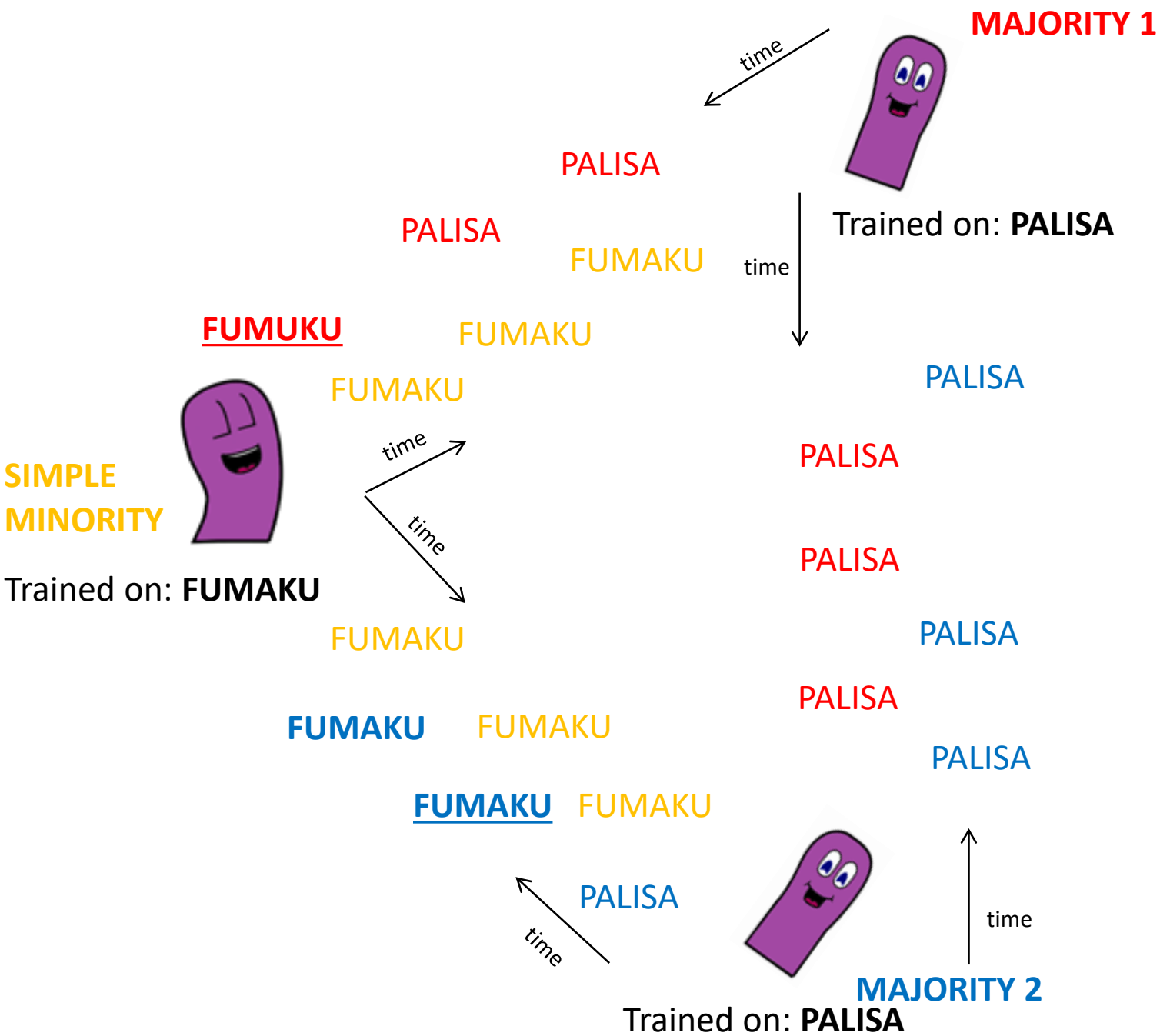
Results

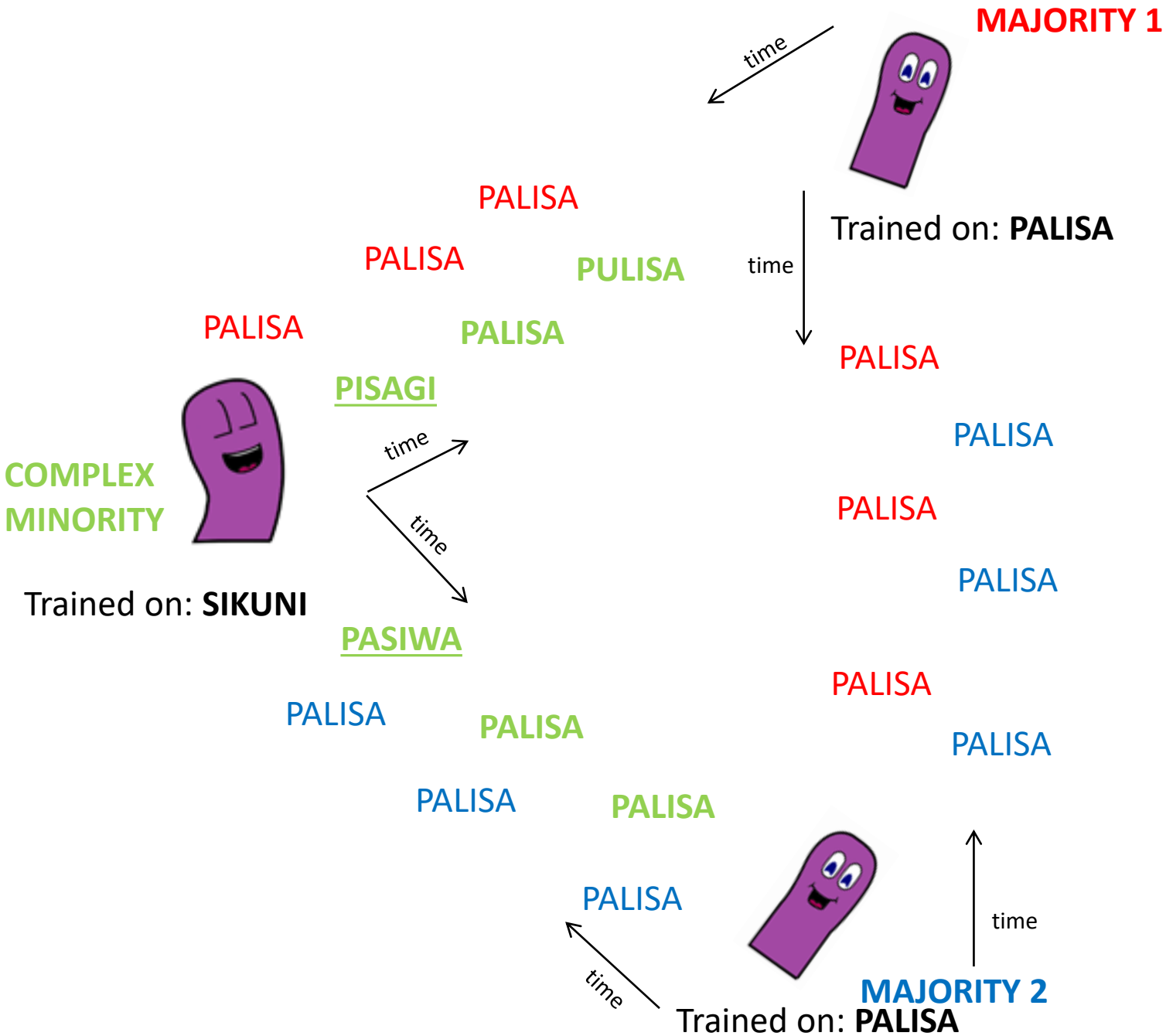
- Learned to a competent level (1 outlier)
- Test items did cause problems at first



Solving the communicative problem

- A robust trend:
 - Complex minority accommodated to the majority
 - Simple minority did not
 - They reached either a mutual understanding, or sometimes also the majority accommodated to the minority





For the 2nd test item

Change

Complex minority – Complex majority

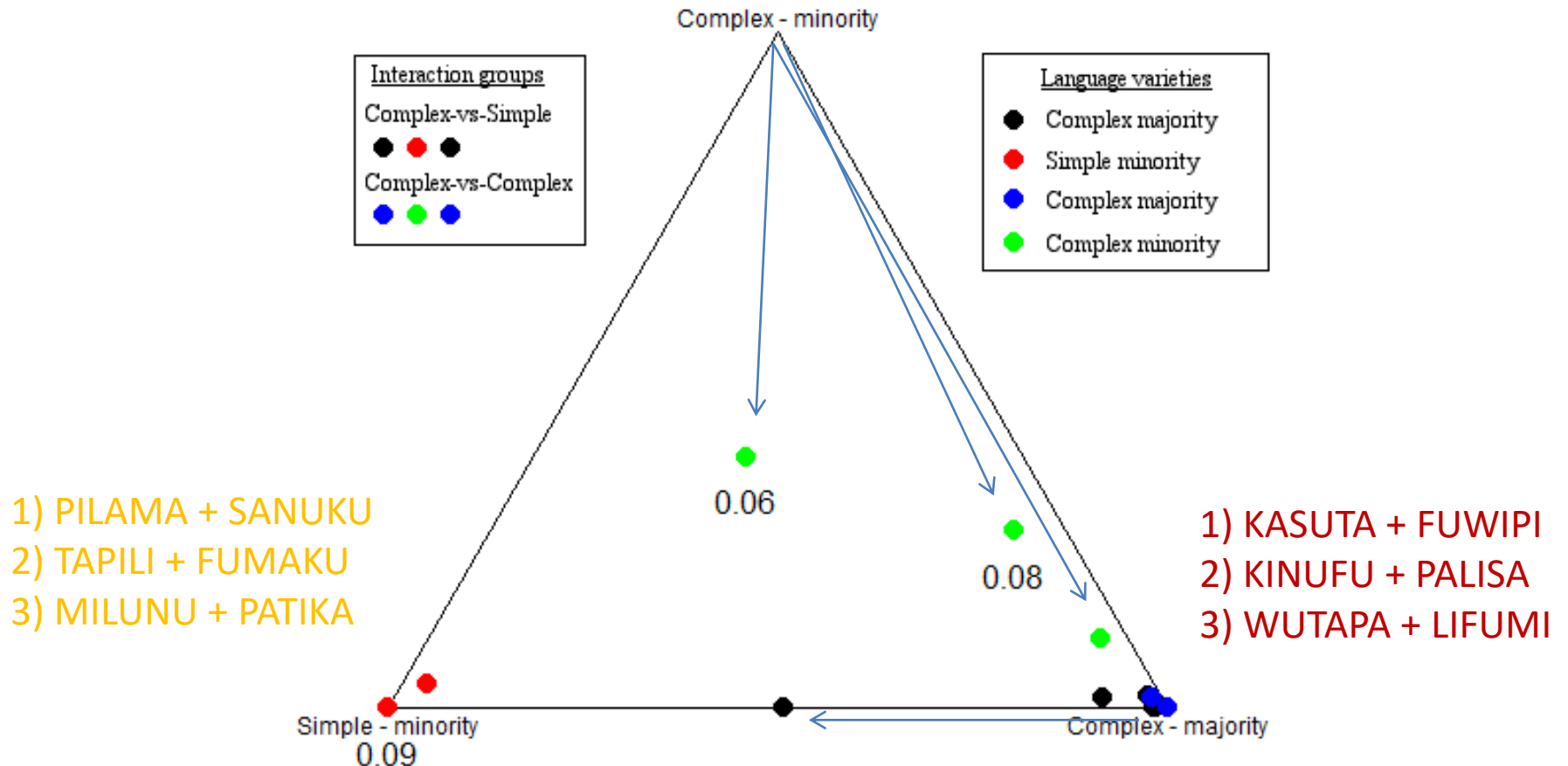
WULAMA	WULAMA	<u>KINUFU</u>	<u>KUSALI</u>	<u>KINUFU</u>	<u>KINUFU</u>
KINUFU	KINUFU	KINUFU	KINUFU	KINUFU	KINUFU
KINUFU	KINUFU	KINUFU	KINUFU	KINUFU	KINUFU

Simple minority – Complex majority

TAPIWI	TAPIWI	TAPIWI	TAPIWI	TAPIWI	TAPIWI
KINUFU	KINUFU	<u>TAPIWI</u>	<u>TAPILI</u>	KINUFU	<u>TAPILI</u>
KINUFU	KINUFU	KINUFU	KINUFU	KINUFU	KINUFU

Turned bilingual

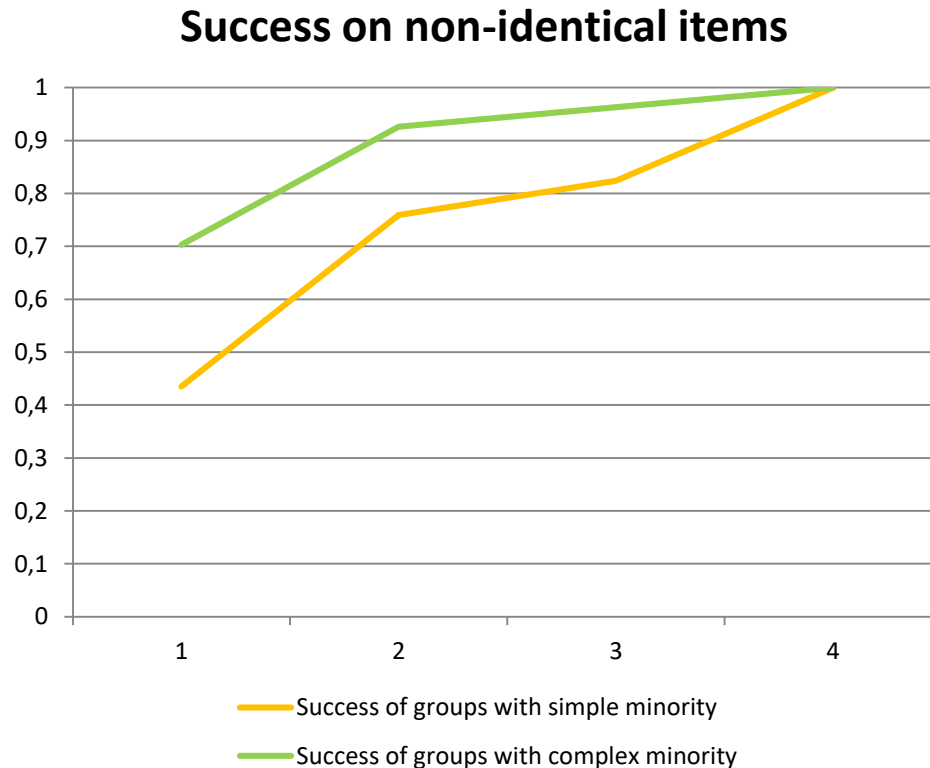
- 1) TUFINU + MILAWA
- 2) WULAMA + SIKUNI
- 3) KANIFI + SULAWU



Distance from training stimuli for the minority, between two conditions
one-tailed pairwise t-test, $t(2) = -13.503 < .01^{**}$

An interesting note

- The simple minority groups lost the game



Discussion 1

- Languages proved learnable within the time-constraints, and the interaction stayed very close to conventions.
- Differences between varieties did cause a communicative difficulty in the beginning.
- All groups reached maximum success on the test items by last round, although the stability of the community conventions is unknown.

Discussion 2

- Current sample size is too small to make strong claims, but a trend is apparent. Further study is needed to confirm/disconfirm.
- The trend exhibits a greater stability for more systematic varieties in contact situations.
- This would allow the threshold of L2-learner innovations needed for simplification to be reduced, given the model.
- The strategic inferiority of maintaining a minority variety is interesting, but a separate topic.

Conclusions

- Placing participants with prelearned miniature languages in contact situations seems to be a viable research paradigm.
- If so, a small range of research questions regarding diffusion of forms and formation of varieties within groups may become better approachable.
- The use of artificial languages should allow the setup to be extended to other structural features of interest.
- Can not make strong conclusions about the results of the current test case just yet.

References 1

- Beckner, C.; Blythe, R.; Bybee, J.; Christiansen, M.; Croft, W.; Ellis N.; Holland, J.; Ke, J.; Larsen-Freeman, D.; Schoenemann, T. 2009. Language is a Complex Adaptive System: Position Paper. *Language Learning*, 59(1): 1-26
- Culbertson, J., & Smolensky, P. (2009). Testing Greenberg's Universal 18 using an artificial language learning paradigm. *Proceedings of NELS*, 1–12.
- Dahl, Ö. (2004). *The growth and maintenance of linguistic complexity*. Amsterdam: John Benjamins
- Fay N., Garrod S., Roberts L. (2008). The fitness and functionality of culturally evolved communication systems. *Philosophical Transactions of the Royal Society: Biological Sciences*. 363:3553–3561
- Galantucci, B., & Garrod, S. (2009). Experimental semiotics: a review. *Frontiers in human neuroscience*.
- Kalnins, D. (2010). *Iterated Learning and Communicative Use of Spoken Language*. (Unpublished master's thesis). University of Edinburgh, Edinburgh.
- Kirby, S., Cornish, H., and Smith, K. (2008). Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proc. Natl. Acad. Sci. U.S.A.* 105, 10681–10686.
- Kusters, W. 2003. *Linguistic Complexity: the Influence of Social Change on Verbal Inflection*. PhD Dissertation, University of Leiden.
- Lupyan G, Dale R. (2010). Language Structure Is Partly Determined by Social Structure. *PLoS ONE* 5(1): e8559

References 2

- McWhorter, J. (2007). *Language Interrupted*. Oxford: Oxford University Press.
- Roberts, G. (2010). An experimental study of social selection and frequency of interaction in linguistic diversity. *Interaction studies*. doi:11.1075/is.11.1.06rob
- Scott-Phillips, T. C., & Kirby, S. (2010). Language evolution in the laboratory. *Trends in cognitive sciences*, 14(9), 411–7. doi:10.1016/j.tics.2010.06.006
- Smith, A.D.M., Skarabela, B. and Tamariz, M. (2010). Exploring the nature of a systematicity bias: an experimental study. In A.D.M. Smith, M. Schouwstra, B. de Boer & K. Smith (Eds.) *The Evolution of Language (EVOLANG 8)* (pp. 289–296). Singapore:World Scientific.
- Smith, K., & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. *Cognition*, 116(3), 444–9.
- Tamariz, M., Cornish, H., Roberts, S. & Kirby, S. (2012). How generation turnover and interlocutor negotiation affect language evolution. In T. Scott-Phillips, M. Tamariz & E. Cartmill, (Eds.) *The Evolution of Language*. Proceedings of the 9th International Conference on the Evolution of Language.
- Tily, H., Frank, M., Jaeger, T. F. (2011). The learnability of constructed languages reflects typological patterns. *Proceedings of CogSci*, 1364–1369.
- Trudgill, P. (2012). On the sociolinguistic typology of linguistic complexity loss. In F. Seifart, G. Haig, N. P. Himmelmann, D. Jung, A. Margetts, and P. Trilsbeek (Eds.) *Potentials of Language Documentation: Methods, Analyses, and Utilization* (pp. 96-104). Honolulu: University of Hawai'i Press

- Thank you!

peeter.tinits@gmail.com



THE UNIVERSITY *of* EDINBURGH



UNIVERSITY *of* TARTU