THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Exploring the Nature of a Systematicity bias: An experimental study

Citation for published version:
Smith, A, Skarabela, B & Tamariz, M 2010, Exploring the Nature of a Systematicity bias: An experimental study. in ADM Smith, B de. Boer, M Schouwstra & K Smith (eds), Proceedings of the 8th International Conference on the Evolution of Language. World Scientific Press, pp. 289-296, The Evolution of Language - 8th International Conference (EVOLANG8), Utrecht, Netherlands, United Kingdom, 14/04/10.

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Proceedings of the 8th International Conference on the Evolution of Language

Publisher Rights Statement:
© Smith, A., Skarabela, B., & Tamariz, M. (2010). Exploring the Nature of a Systematicity bias: An experimental study. In A. D. M. Smith, B. de. Boer, M. Schouwstra, & K. Smith (Eds.), Proceedings of the 8th International Conference on the Evolution of Language. (pp. 289-296). World Scientific Press.

OPEN ACCESS

EXPLORING THE NATURE OF A SYSTEMATICITY BIAS: AN EXPERIMENTAL STUDY

ANDREW D. M. SMITH, BARBORA SKARABELA & MÓNICA TAMARIZ

Language Evolution and Computation Research Unit,
Linguistics and English Language,
The University of Edinburgh,
Dugald Stewart Building, 3 Charles Street, Edinburgh, EH8 9AD
{andrew;barbora;monica}@ling.ed.ac.uk

In this study, we tested the circumstances under which cultural evolution might lead to regularisation, even in the absence of an explicit learning bottleneck. We used an artificial language experiment to evaluate the degree of structure preservation and the extent of a bias for regularisation during learning, using languages which differed both in their initial levels of regularity and their frequency distributions. The differential reproduction of regular and irregular linguistic items, which may signal the existence of a systematicity bias, is apparent only in languages with skewed distributions: in uniformly distributed languages, reproduction fidelity is high in all cases. Regularisation does happen despite the lack of an explicit bottleneck, and is most significant in infrequent items from an otherwise highly regular language.

1. Introduction

One of the most striking features of language is its extensive use of conventionalised systematicity in representing and conveying meaning. Signal-meaning systematicity can be found at many levels of linguistic analysis, from simple morphological paradigms to complex syntactic and discourse constructions. The actual signal-meaning mappings in a particular language depend, as in any adaptive system, on the language's initial conditions (Gell-Mann, 1994), and the characteristics of its usage history, such as regularities which may have appeared by chance (Wray, 2002), the order in which items are learnt (Foss, 1968; Schyns & Rodet, 1997), or the way in which language use is negotiated (Galantucci, 2005). Researchers have proposed several different sources for linguistic systematicity, including cognitive biases (Pinker, 1999), the nature of the social and communicative context in which language is used (Wray & Grace, 2007), the frequency of the input (Bybee & Hopper, 2001), or a transmission bottleneck (Kirby, 2001).

A number of models, both computational (Brighton, Smith, & Kirby, 2005) and experimental (Kirby, Cornish, & Smith, 2008), have shown how a bottleneck on the transmission of linguistic items leads to the emergence of regularity and compositionality. It has, furthermore, been proposed as the mechanism which

may have led to the emergence of compositional language from a holistic protolanguage (Kirby, 2002). It is not clear, however, that an explicit bottleneck is indeed a necessary and sufficient condition for regularity to arise, and so we present a study exploring the effects of frequency on the development of regularity in linguistic structure in the absence of an explicit transmission bottleneck.

We use an artificial language experiment with multiple languages which differ in the degree of regularity in their signal-meaning correspondences. Although artificial language systems have a long pedigree in psycholinguistic research, dating back to Esper (1925) and Wolfle (1933), it is only in recent years that they have begun to be used to investigate the adaptation and evolution of linguistic structure (Kirby et al., 2008; Beqa, Kirby, & Hurford, 2008; Tamariz & Smith, 2008). This study differs from existing artificial language experiments in that it studies the regularisation of internal linguistic structure, rather than the regularisation of inconsistent, unpredictable input (Hudson Kam & Newport, 2005; Wonnacott & Newport, 2005). We explicitly investigate the impact of the degree of regularity in the initial language (see also Tamariz & Smith, 2008) and the effects of different frequency distributions on language reproduction and regularisation, rather than using only languages with a uniform frequency distribution.

We focus on an exploration of how the *regularity of a language* and the *frequency distribution* of its items affect: (i) the fidelity with which the original form-meaning associations are reproduced; (ii) and the regularisation of irregular items. We predicted that languages would be reproduced better as their overall regularity increased; that within each language, regular items would be reproduced better than irregulars, and frequent items would be reproduced better than infrequents; and that if regularisation occurred, it would be most likely in infrequent items.

2. Method

Participants 44 monolingual English native speakers took part in the experiment, with four being excluded from the analysis (three due to software failure and one who failed to comply with the requirements of the study); each was paid £4.

Materials Participants were asked to learn pairings of pictures depicting familiar activities and nonsense words. Twelve different activities were chosen, using pictures designed to be culturally and linguistically neutral, and used by foreign language instructors; these pictures were then modified (by adding hair, clothing etc.) to produce two different versions: male and female[a]. Nonsense words were created according to English phonotactics, all containing three syllables (a bisyllabic root and a monosyllabic suffix); each language contained equal num-

---

[a]Original pictures are freely available at `http://tell.fll.purdue.edu/JapanProj/FLClipart/`; modified stimuli used in the experiment at `http://www.ling.ed.ac.uk/~andrew/systematicity/gender_activity_stimuli.html`.

bers of roots and suffixes with a variety of syllable structures (roots: VCV, VCCV, CVCV, CVCCV, CCVCV, CCVCCV; suffixes: CV; CVC). The experiment was developed using Psyscope X and run on laptops with Mac OS X.

Design Five types of experimental language were created, all sharing the same 24 meanings (12 activities x 2 genders), but having different words associated with these meanings. Two languages were created for each language type, differing in the forms of the roots and suffixes, but structurally identical. The language types differed in their level of regularity, or the degree to which there is a systematic correspondence between forms and their meanings. We measured regularity in signal-meaning correspondences along two dimensions: a root corresponding to the activity (e.g. swimming, drinking); and a suffix corresponding to the natural gender of the agent of the activity (male, female):

0% regular: The words for all twelve activities were compiled from 24 roots and 24 suffixes[b]. This language type is effectively fully holistic, with an independent phonological form for each meaning.

33% regular: Four activities had regular mappings (i.e. four roots and two suffixes for the eight meanings), while the other eight activities were compiled from 16 roots and 16 suffixes.

50% regular: Six activities had regular mappings, the other six were compiled from 12 roots and 12 suffixes.

67% regular: Eight activities had regular mappings, the other four were compiled from 8 roots and 8 suffixes.

100% regular: All twelve activities had regular mappings (i.e. twelve roots and two suffixes). This language type is fully compositional.

Each language was presented in two different distributions: with either a uniform distribution (all items were seen equally frequently) or a skewed distribution (half the items were seen nine times more frequently than the other half).

Procedure Participants were explicitly told that they would see pictures on the computer screen, with descriptions, and that they would subsequently be asked to remember which description went with which picture. Participants were seated in front of a computer screen in a quiet room on the university premises. Training consisted of exposure to the whole language via 240 exposures: at each exposure, the picture was presented first, followed 500ms later by a description below it (see

---

[b]Details of the forms in all the languages can be found at `http://www.ling.ed.ac.uk/~andrew/systematicity/expt1_languages.html`.

Fig. 1). In testing, individual pictures appeared on the screen, and participants were asked to type descriptions as they remembered. After testing, they completed a questionnaire and were debriefed on the aims of the experiment. Training and testing occurred over one self-paced session, which lasted on average 25 minutes.
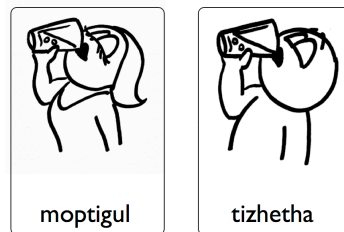

Figure 1.   Examples of two separate training exposures: each contains a picture of an activity paired with a nonsense word. Meanings are differentiated by the gender of the activity's agent. Words consist of a bisyllabic root (here: **mopti-, tizhe-**) and a monosyllabic suffix (here: **-gul, -tha**).

3. Results

3.1. *Reproduction Fidelity*

Figure 2 shows the reproduction fidelity of suffixes by initial language regularity. We ran a series of ANOVAs with regularity and frequency distribution of language as the independent variables, and found a significant effect of both variables on the reproduction fidelity of suffixes (frequency distribution: $F_{(1,38)} = 4.63, p = .04$; regularity: $F_{(4,35)} = 2.86, p = .04$), but a significant effect only of frequency distribution on the reproduction of roots (frequency distribution: $F_{(1,38)} = 10.29, p < .01$; regularity: $F_{(4,35)} = 0.50, p = .74$). Overall, suffixes are significantly better reproduced as the regularity of the language increases, and in uniformly distributed languages rather than in skewed languages. Roots are reproduced better in uniformly distributed languages, but equally well in languages with different levels of regularity. This latter finding is likely to be due to the nature of regularity in these artificial languages: if there is any regularity in the languages it tends to concern suffixes rather than roots, because a regular root occurs across only two lexical items, whereas a regular suffix can occur with up to 12 different roots.

Considering the frequency distributions separately, we found no significant effects on the reproduction fidelity of different forms in uniform languages: regular and irregular suffixes were reproduced equally well ($F_{(1,30)} = 1.13, p = .30$), as were regular and irregular roots ($F_{(1,30)} = 0.14, p = .71$). Regular forms
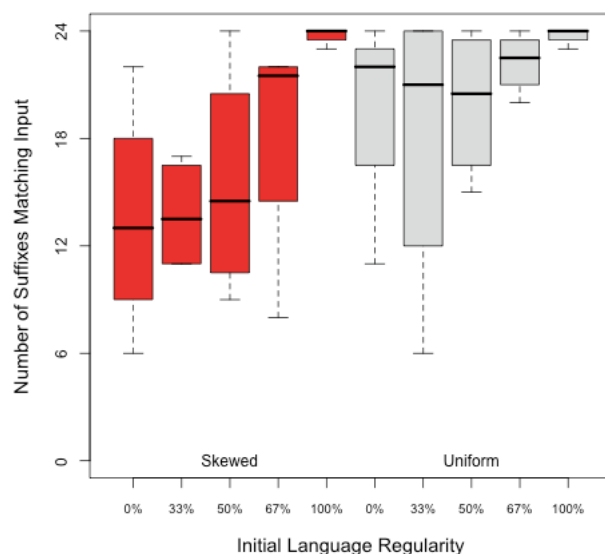
Figure 2. Mean reproduction fidelity of suffixes, by frequency distribution and language regularity.

were reproduced equally well across the four language types[c] (suffixes: $F_{(3,12)} = 0.64, p = .61$; roots: $F_{(3,12)} = 1.05, p = .41$), as were irregular forms (suffixes: $F_{(3,12)} = 0.35, p = .79$; roots: $F_{(3,12)} = 0.33, p = .80$).

In skewed languages, however, regular suffixes were reproduced significantly better than irregular ones ($F_{(1,62)} = 10.12, p < .001$), and frequent suffixes were reproduced significantly better than infrequent ones ($F_{(1,62)} = 19.33, p < .001$). There is an additional interaction between item regularity and item frequency: the reproduction fidelity of infrequent irregulars is even lower than would be expected by combining the individual effects of regularity and frequency ($F_{(1,60)} = 4.53, p = .04$). The same is not true, however, for roots: although frequent roots are reproduced better than infrequent roots ($F_{(1,62)} = 35.86, p < .001$), regular and irregular roots are reproduced equally well ($F_{(1,62)} = 0.45, p = .51$).

3.2. *Regularisation*

In the cases in which reproduction is not faithful (such as the irregular infrequent items in the skewed languages), therefore, can we discern any trends towards reg-

---

[c]There are only four language types in these analyses because the 0% languages have no regular forms; likewise, 100% languages have no irregulars.

ularisation which might be regarded as a systematicity bias? Is there a tendency for participants to regularise forms which were originally irregular, and if so, what role do language regularity and frequency distributions play?

To examine what happens in the instances when originally irregular suffixes were not faithfully reproduced, we focused on the irregular items in the input languages. Regularisation was operationally defined as a change from an originally irregular suffix, which was therefore used on only one item in the input language, to a suffix which was used on multiple items in the output language to convey the *same* gender (i.e. either male or female). Zero suffixes were excluded, so deletion of the suffix across multiple items was not counted as regularisation in our analysis. The regularised suffix could therefore be either: (i) an originally regular suffix whose range has increased to encompass one or more irregular items; (ii) an originally irregular suffix whose range has increased to encompass one or more additional items; (iii) a novel suffix used on two or more items. According to these criteria, 8.75% of the originally irregular items were regularised (42 items in total: 9 originally regular suffixes; 20 originally irregular suffixes; 13 novel suffixes).
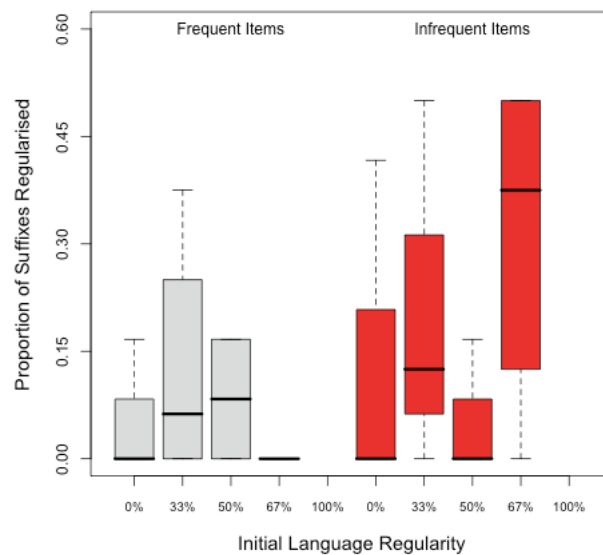

Figure 3.   Regularisation of suffixes in languages with skewed frequency distributions, by item frequency and language regularity.

We ran a series of ANOVAs with language regularity and frequency dis-

tribution as the independent variables. While we predicted that regularisation would be more likely to take place in the skewed rather than the uniform languages, we did not find a significant difference between the two types of languages ($F_{(1,46)} = 0.91, p = .35$), nor did we find a significant trend towards regularisation with increasing language regularity ($F_{(3,44)} = 0.69, p = .56$). We further compared suffix regularisation in the languages with uniform frequency distributions, and found no significant difference between the process in languages with different degrees of regularity ($F_{(3,12)} = 1.14, p = .37$).

Figure 3 shows details of regularisation in the skewed languages. Although language regularity did not yield a significant effect overall ($F_{(3,28)} = 0.72, p = .55$), infrequent suffixes were marginally significantly more likely to undergo regularisation than frequent suffixes ($F_{(1,31)} = 2.91, p = .10$). Looking at this effect in more detail, we found that the difference in regularisation between frequent and infrequent irregulars is significant only in the language with the highest level of regularity[d] (67%, $F_{(1,6)} = 6.82, p = .04$). This is indicative of an interaction between language regularity and frequency, which may be obscured by the relatively low number of participants in this study.

4. Summary and Conclusion

The potential systematicity bias, which would be signalled by differences in reproduction fidelity between regular and irregular items, is more likely to be observed in languages with skewed frequency distributions rather than in languages where items occur with equal frequency, because uniform languages are better learnt. Regularisation is also affected by both frequency and the existing level of language regularity: although there is no difference between frequent and infrequent items from languages with relatively low levels of overall regularity, in languages with high regularity, infrequent irregulars are regularised while frequent regulars are reproduced more faithfully.

These results suggest that the skewed frequency distribution reveals the systematicity bias by acting as an *implicit* bottleneck, which makes it very difficult for language users to reproduce infrequent items accurately, and therefore leads to regularisation even in the absence of an explicit bottleneck. The meanings humans communicate about in social interactions are notably skewed in their frequency distribution — we talk about a few favourite things most of the time, and about many other things much less often. This study shows that such a distribution may have played a role in the evolution of language by promoting regularisation.

Acknowledgements

---

[d]Note that regularisation cannot apply to fully regular (100%) languages.

References

Beqa, A., Kirby, S., & Hurford, J. R. (2008). Regular morphology as a cultural adaptation: Non-uniform frequency in an experimental iterated learning model. In A. D. M. Smith, K. Smith, & R. Ferrer i Cancho (Eds.), *The Evolution of Language (EVOLANG 7)* (pp. 401–402). World Scientific.

Brighton, H., Smith, K., & Kirby, S. (2005). Language as an evolutionary system. *Physics of Life Reviews*, *2*(3), 177-226.

Bybee, J. L., & Hopper, P. J. (Eds.). (2001). *Frequency and the emergence of linguistic structure.* Amsterdam: John Benjamins.

Esper, E. A. (1925). *A technique for the experimental investigation of associative interference in artificial lingustic material.* Philadelphia: Linguistic Society of America.

Foss, D. J. (1968). Learning and discovery in acquisition of structured material – effects of number of items and their sequence. *Journal of Experimental Psychology*, *77*(2), 341–344.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, *29*, 737-767.

Gell-Mann, M. (1994). *The quark and the jaguar.* New York: Freeman.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: the roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*(2), 151-195.

Kirby, S. (2001). Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. *IEEE Journal of Evolutionary Computation*, *5(2)*, 102–110.

Kirby, S. (2002). Learning, bottlenecks and the evolution of recursive syntax. In E. Briscoe (Ed.), *Linguistic evolution through language acquisition: Formal and computational models* (pp. 173–203). Cambridge University Press.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the lab: an experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681–10686.

Pinker, S. (1999). *Words and rules.* Weidenfeld & Nicolson.

Schyns, P. G., & Rodet, L. (1997). Categorization creates functional features. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*(3), 681–696.

Tamariz, M., & Smith, A. D. M. (2008). Regularity in mappings between signals and meanings. In A. D. M. Smith, K. Smith, & R. Ferrer i Cancho (Eds.), *The Evolution of Language (EVOLANG 7)* (pp. 315–322). World Scientific.

Wolfle, D. L. (1933). The relative stability of first and second syllables in an artificial language. *Language*, *9*(4), 313-315.

Wonnacott, E., & Newport, E. L. (2005). Novelty and regularization: The effect of novel instances on rule formation. In A. Brugos, M. R. Clark-Cotton, & S. Ha (Eds.), *BUCLD29: Proceedings of the 29th Annual Boston Conference on Language Development.* Somerville, MA: Cascadilla Press.

Wray, A. (2002). *Formulaic language and the lexicon.* Cambridge: CUP.

Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, *117*(3), 543-578.