University Press Potsdam

Article published in:

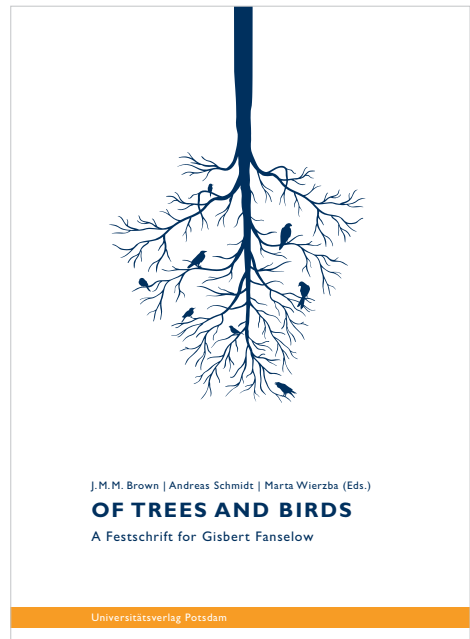J.M.M. Brown | Andreas Schmidt | Marta Wierzba (Eds.)

OF TREES AND BIRDS

A Festschrift for Gisbert Fanselow

Universitätsverlag Potsdam

Is it a bird? Is it a mammal? Perspectives on the learnability / trainability of new grammatical constructions

Ina Bornkessel-Schlesewsky, Matthias Schlesewsky

School of Psychology, Social Work and Social Policy
University of South Australia

1 Introduction: strange animals, strange constructions – an observer's perspective

When the second governor of the colony of New South Wales, Captain John Hunter, sent drawings and a skin of a platypus to England in 1797, they caused quite a stir among European naturalists (Hall 1999). The unusual features of the animal – a duck-like beak combined with a body resembling that of a mole, and hence a mammal – defied classification within the biological taxonomies of the late 18th and early 19th centuries. Its combination of characteristics appeared so paradoxical even before the discovery that the platypus also lays eggs and succles its young, that some believed it to be a hoax: a fabricated specimen with a duck's bill sewn onto the body of a mole.[1]

The story of the European discovery of the platypus is a potent example of how difficult it can be to understand and accept exemplars that fall

---

[1]. For a humorous account of this and other occurrences in Australian history, see Hunt (2013).

outside the entrenched expectations of one's normal classificatory system. It illustrates a phenomenon with which we are also confronted in linguistics and which has been a continuing research interest of Gisbert Fanselow's: the question of how speakers deal with constructions outside of their own grammatical inventory and the conditions under which they are able to learn and integrate these into that existing inventory.

In an initial foray into this area, Fanselow et al. (2008) observed something quite extraordinary (though, arguably, not as extraordinary as the platypus must have appeared to late 18th century European naturalists). By simply exposing Northern German speakers to sentence constructions that are unavailable in their native dialectal variety but that are possible, for example, in some Southern German (e.g. Bavarian) varieties, Fanselow and colleagues were able to (implicitly) train them to use these constructions within a matter of 6 to 8 weeks. Strikingly, learning of the complex constructions in question (long-distance *wh*-extractions out of that clauses) took place via only two sessions of exposure and resulted in indistinguishable usage patterns to a Bavarian control group as measured via a sentence completion questionnaire.

The authors' proposed explanation for this increasing acceptance of "non-native" constructions and the concomitant increase in their usage was that exposure-based training is possible for structures that are compatible with one's overall grammatical system. In other words: *wh*-extraction out of *that*-clauses is assumed to be possible as part of the grammatical system underlying all German varieties, with regional variation in the scope of extractability being due to extragrammatical factors.

The ease of trainability for *wh*-extractions contrasts with another phenomenon that, in the intuition of the current authors, is considerably more difficult to train: dialectal variation in auxiliary selection in German (for empirical data, see Keller & Sorace 2003). Intransitive "position" verbs combine with *haben* ('to have') in Northern German varieties, but with *sein* ('to be') in Southern German varieties (e.g. *Der Kakapo* hat/ist *vor der Kamera gesessen*; word-by-word translation: 'The kakapo has/was in-front-of the camera sat'; 'The kakapo sat in front of the camera'). While speakers of both Northern and Southern German varieties are regularly exposed to both variants through the media and literature,

276

it seems to us that this does not influence usage patterns. This phenomenon clearly differs from the *wh*-extractions discussed above in that it involves regional alternatives to express similar meanings, while an extended scope for *wh*-extractions enriches the construction inventory for Northern German speakers.

Nevertheless, the contrast between the two phenomena raises a number of questions. Where are the boundaries of trainability? What is the best definition of a "grammatical system" and under which conditions is a system or are parts of a system susceptible to training effects? Do acceptance of a construction and its usage differ as a function of exposure or do they always go hand in hand?

In this chapter we will touch upon these questions, which we believe are also still highly pertinent to Gisbert Fanselow's current research interests. Moreso, however, we would like to discuss three observations from our own work that may be of interest in the context of this line of research. The first two – involving word order and case marking, respectively – stem from language learning studies and touch upon the question of how readily speakers are able to learn constructions that differ in the degree of similarity to their first language (L1). The third, by contrast, returns to the question of how the acceptability of a construction outside of one's native variety or language is influenced by the conditions under which it is encountered, including the social characteristics of the speaker.

2 Phenomenon one: word order

Word order phenomena are the bread and butter not only of syntactians, but also of language typologists, and psycholinguists/neurolinguists interested in sentence comprehension and/or language acquisition. As such, they appear ideally suited to addressing questions of learnability and trainability and how these relate to grammatical systems. In a recent study in our laboratory (Cross et al. in preparation), we examined how L1 English speakers learn new word order regularities by exposing them to a modified miniature version of Mandarin Chinese. We chose Mandarin because it has constructions that are (superficially)

highly similar to English (NP–Verb–NP), as well as others that are quite different in requiring a verb-final word order (NP–coverb–NP–Verb). In addition, the overall system introduces (some) flexibility in the relative order of Actor/Subject and Undergoer/Object arguments[2] that is largely unavailable in English:

- NP–V–NP orders can be either Actor–V–Undergoer or Undergoer–V–Actor

- V-final orders containing coverbs have a fixed argument order that depends on the coverb: Actor–*ba*–Undergoer–V and Undergoer–*bei*–Actor–V. Reversal of Actor and Undergoer is not allowed within these constructions.

Participants (monolingually raised, native speakers of Australian English) first learned the vocabulary of the miniature language via a take-home booklet. They then implicitly learned possible constructions in the lab via pictures depicting transitive events that were accompanied by matching sentences. Following an initial learning session comprising 128 sentences, participants performed a baseline grammaticality judgement task (without pictures), in which they judged whether 288 novel sentences were possible sentences in the miniature language or not. Following a delay period of approximately 9 hours, participants completed a second (delayed) grammaticality judgement task, in order to examine consolidation of the new grammatical knowledge.

For constructions without coverbs, participants showed a reasonably high ($\sim$ 70%) accuracy in judging both types of verb-medial sentences as acceptable, and this did not differ between the baseline and delayed sessions. However, accuracy increased in rejecting verb-final orders without coverbs from baseline to delayed testing (to a mean of approximately 75% at delayed testing for both Actor–Undergoer–V and Undergoer–Actor–V). We take this to suggest that interference with participants' L1 impedes consolidation beyond the baseline level. (There also seems

---

2. Note that we equate Actor and Subject and Undergoer and Object here to make clear that our experiment did not employ constructions suited to differentiating between the two. We do not mean to suggest that the two sets of terms are equivalent.

to be an additional influence of whether participants slept during the delay period or not, which is beyond the scope of this chapter.)

For constructions with coverbs, participants readily learned (again, at a mean level of approximately 70% accuracy) that the basic word order should be verb final, i.e. that the presence of a coverb requires a verb-final order. By contrast, they did not learn the requirements for Actor-Undergoer order for each coverb and thus tended to accept all orders which adhered to the basic NP–coverb–NP–V schema.

In summary, native speakers of English can learn new constraints on verb position quite readily, but have difficulty in learning constructions that don't conform to the strongest (most valid) relational interpretive cue in their L1, namely Actor–before–Undergoer (cf. MacWhinney et al. 1984). Importantly, this experiment revealed an additional dimension of inter-individual differences: statistical learning ability (cf. Daltrozzo et al. 2017) predicted judgement accuracy for grammatical constructions, thus demonstrating that linguistic chracteristics are only one influencing factor in regard to the learnability/trainability of new grammatical constructions.

3 Phenomenon two: case marking

A second phenomenon that we would like to discuss in this regard is case marking. In spite of a wealth of research on different case marking systems, very little is currently known about how L1 speakers of one system (e.g. nominative-accusative) learn another (e.g. ergative-absolutive, let alone differential case marking systems based on animacy distinctions). With another miniature language learning study (Wang et al. in preparation), we thus aimed to examine whether speakers of a case-marking language (here: L1 German) are able to transfer their relational knowledge about case to a system based on different features.

To this end, we constructed two miniature languages (LG1, LG2) based on Hindi, but manipulated the case marking rules to be dependent on animacy. Both languages followed a *distinguishability*-based strategy to mark U arguments (e.g. Comrie 1989), i.e. animate undergoers are marked in order to distinguish them from actor arguments. This is a

common differential case marking pattern found in a range of languages including Hindi and Spanish (Aissen 2003, von Heusinger 2008, Mohanan 1994). For actor arguments, by contrast, the marking strategies differed between the two miniature languages: LG1 employed an *identification* strategy, i.e. mark prominent, animate actors (de Hoop & Malchukov 2008), while LG2 again used a distinguishability strategy, i.e. mark atypical, inanimate actors. An identification-based actor-marking strategy is attested, for example, in Manipuri (Bhat & Ningomba 1997), while Fore is an example of a language that employs a distinguishability-based strategy (Scott 1978).

The learning and testing paradigm employed was very similar to that utilised by Cross et al. (in preparation), with the exception that there was a slightly longer delay interval (up to three days) between the first and second sessions and that the second testing session was also preceded by a second training session. Furthermore, since training was spread across two sessions, participants were only exposed to 60 (grammatical) sentence–picture pairs per session; with testing sessions involving 120 grammatical and ungrammatical sentences. Participants were monolingually raised native speakers of German.

Results revealed two main observations: (a) judgement accuracy improved from session one (mean accuracy: 85.7%) to session two (mean accuracy: 93.2%); and (b) accuracy was generally higher for LG1 (mean across both sessions: 94.8%) versus LG2 (mean: 84.1%). The increase in accuracy from session 1 to session 2 was comparable across both languages and reaction times (RTs) mirrored the accuracy pattern (i.e. faster RTs for the second session and for LG1). Finally, accuracy values showed higher inter-individual variability for LG2 learners (standard deviation for grammaticality judgements across both sessions was 36.6%, as opposed to 22.1% for LG1).

We interpret these findings as follows. Language 1 offers a relatively easy general learning strategy, namely to mark all animate arguments. (Note, however, that participants still needed to learn the correct markers for actors and undergoers and that this involved correct role identification given that the languages both allowed for argument order variability.) Matters are more complex for LG2, in that participants needed to learn to mark deviations from prototypical transitive constructions in

the sense of Comrie (1989), i.e. from constructions with an animate actor and an inanimate undergoer. Given the complexity of this task and the comparably small number of training sentences, it is quite striking, in our view, that participants were nevertheless able to attain a very respectable level of accuracy for LG2. All in all, then, native speakers of German appear to be able to pick up a new case marking pattern relatively easily, possibly due to the fact that the animacy-based marking system employed here was sufficiently distinct from the German case system.

These observations lead us to several new hypotheses: (a) this case-system learning task should be more difficult for speakers of non-case-marking languages (a hypothesis currently being tested by Luming Wang and colleagues); (b) the task may have been harder for German L1 speakers had a "plain" ergative pattern been employed rather than a differential case marking pattern based entirely on animacy. The reasoning for (b) is that learning an ergative system requires speakers with an accusative L1 to switch alignment patterns from S/A versus P to S/P versus A, which should result in a higher degree of interference from the L1 than the animacy-based case-marking rules employed here. Initial evidence for this claim stems from a previous study of ours on the processing of case marking and aspect in Hindi Choudhary et al. (2009), in which we observed an unexpected acceptability asymmetry for aspect-case mismatches. Specifically, while grammatical descriptions of Hindi (e.g. Mohanan 1994) state that A arguments should be marked with ergative in perfective constructions but remain unmarked (i.e. absolutive/nominative) in imperfective constructions, mean acceptability ratings for sentences containing ungrammatical nominative marking in a perfective context yielded a mean acceptability of 31%, compared to just 8% for ungrammatical ergative marking in an imperfective context. We speculated that this asymmetry may reflect the exposure of native Hindi speakers to speakers with other (non-ergative) language backgrounds. As (at least some of) these speakers have difficulty with ergative case marking, they tend to use nominative where ergative is required, thus rendering nominative-A in a perfective context an attested construction, while speakers virtually never encounter ergative-A in an imperfective context. In the next section, we elaborate on the potential

consequences of this finding for the training conditions under which individuals most readily learn new grammatical constructions.

4 Optimal training conditions?

Choudhary et al. (2009)'s observation of a higher acceptability for supposedly ungrammatical constructions in Hindi which are used by non-native speakers without an ergative language background is in line with results by Hanulíková et al. (2012). These authors found qualitatively different electrophysiological brain responses in native speakers of Dutch for ungrammatical Dutch sentences that were recorded by a speaker with a foreign accent (in comparison to a native speaker). Results such as these attest to a hearer's ability to adapt their grammatical expectations to their assumptions about the speaker's language background. This suggests that the acceptability of structures that are normally deemed unacceptable within a linguistic in-group should increase when they are produced by speakers who can clearly be attributed to the region/language background that is congruent with the construction in question.

With this, we return to the regional variation in auxiliary selection that was discussed at the beginning of this chapter. To examine whether the acceptability of *haben/sein* selection would vary depending on the provenance of the person uttering the sentence in question, we asked participants from Northern versus Southern Germany to listen to sentences that were recorded by a Northern and a Southern speaker, respectively (Philipp et al. in preparation). In order to emphasise the regional background of the speakers, the experiment commenced with a short (fictional) recording, in which each speaker introduced themselves and where they were from. For the regionally varying position verbs (see above), we observed an intriguing – and unexpected – pattern. Firstly, there was a general difference between participants from Nothern and Southern Germany, which resembles the results from Keller & Sorace (2003): while Northern participants showed a clear preference for *haben* versus *sein*, Southern participants judged both variants to be equally, but only modestly acceptable (mean acceptability ratings between 63% and 73%). Secondly – and this was the unexpected result – neither group

showed an influence of speaker provenance on their acceptability ratings. Thus, while participants' own linguistic experience and thus their exposure to these constructions clearly influenced their acceptability judgements, they appeared to "ignore" their explicit knowledge about the distribution of the two auxiliaries in relation to a speaker's regional origin.

This could reflect a number of different influences, which we can't disentangle on the basis of the current results. It is, of course, entirely possible that the preference for *sein* in Southern listeners is weaker than the preference for *haben* in Northern listeners, as already suggested by Keller & Sorace (2003)'s results. However, to demonstrate this, one would need to examine whether there is an equivalent asymmetry in production – which we doubt, though this has not, to the best of our knowledge, been investigated in conjunction with acceptability ratings to date. Stipulating, then, that Southern German speakers have a preference for the production of *sein* but no strong preference in their acceptability judgements (in contrast to Northern German speakers), this suggests a political influence of perceived language prestige, with Northern dialects typically (and, of course, arbitrarily from a linguistic perspective) referred to as "standard" or "high" German. A similar conclusion might be drawn for Choudhary et al. (2009)'s results on Hindi.

All in all, these results highlight the multiplicity of influences on potential training scenarios. On the one hand, there is evidence to suggest that speaker identity plays a role in determining the acceptability of (normally) unacceptable structures (Hanulíková et al. 2012). On the other, our results on auxiliary selection in German indicate that potential speaker-based influences cannot be examined independently of the socio-linguistic circumstances in which the grammatical variability in question occurs.

5 Final remarks

Fanselow et al. (2008)'s suggestion that we may be able to push the boundaries of a grammatical system via exposure to new constructions remains extremely intriguing. Inspired by this proposition, we have

examined some of the potential complexities involved in testing out these boundaries. From our perspective, the ability for exposure and training to increase acceptability appears to transcend varieties of the same language system and, in certain circumstances, even be hindered by interference from one's current grammatical system (cf. the notion of "transfer" in L2 learning, e.g. MacWhinney 1992). However, this of course concerns the learning of a new language and is thus not fully comparable to the situation examined by Fanselow et al. (2008). Nevertheless, we have also observed certain tendencies towards rigidity within one language system, as attested to by the auxiliary selection data from German. In combination with approaches to inter-individual variability (recall the results by Cross et al. and correlations with statistical learning ability), these types of studies open up an entirely new field.

We suggest that learning and training paradigms present a fruitful approach to understanding the capacity of language systems – and their various subcomponents – to be modulated by exposure. Structured exposure to different groups with different language usage patterns could be an interesting new addition to this type of learning paradigm, allowing it to be adapted to study training effects within an individual's L1 as opposed to the learning of a completely new system. This type of approach could be extended to examining language change, by allowing us to observe language systems *in statu nascendi*. While we will not be able to explain in this way how innovation arises, such paradigms may be able to provide insights into the conditions under which innovations spread and thrive.

We would like to conclude by thanking our mentor, Gisbert Fanselow, for inspiring us to work on fascinating topics such as these, for always being open to new ideas and approaches and for his selfless support. Happy birthday, Gisbert, and all the very best for many happy and productive years to come!

Bibliography

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21(3). 435–483.

Bhat, D. N. Shankara & M. S. Ningomba. 1997. *Manipuri grammar.* München: Lincom Europa.

Choudhary, Kamal Kumar, Matthias Schlesewsky, Dietmar Roehm & Ina Bornkessel-Schlesewsky. 2009. The N400 as a correlate of interpretively-relevant linguistic rules: Evidence from Hindi. *Neuropsychologia* 47(13). 3012–3022.

Comrie, Bernard. 1989. *Linguistic universals and language typology.* Oxford, England: Blackwell.

Cross, Zachariah R., L. Zou-Williams, Mark J. Kohler, R. F. Helfrich, S. Coussens, M. G. Gaskell, Matthias Schlesewsky, R. T. Knight & Ina Bornkessel-Schlesewsky. In preparation. Sleep-facilitated consolidation of higher-order language combinatorics: Insights from oscillatory neuronal activity.

Daltrozzo, Jerome, Samantha N. Emerson, Joanne Deocampo, Sonia Singh, Marjorie Freggens, Lee Branum-Martin & Christopher M. Conway. 2017. Visual statistical learning is related to natural language ability in adults: An ERP study. *Brain and Language* 166. 40–51.

Fanselow, Gisbert, Reinhold Kliegl & Matthias Schlesewsky. 2008. Syntactic variation in German *wh*-questions: Empirical investigations of weak crossover violations and long *wh*-movement. *Linguistic Variation Yearbook* 5. 37–63.

Hall, Brian K. 1999. The paradoxical platypus. *BioScience* 49(3). 211–218.

Hanulíková, Adriana, Petra M. van Alphen, Merel M. van Goch & Andrea Weber. 2012. When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience* 24(4). 878–887.

von Heusinger, Klaus. 2008. Verbal semantics and the diachronic development of DOM in Spanish. *Probus* 20. 1–31.

de Hoop, Helen & Andrej L. Malchukov. 2008. Case-marking strategies. *Linguistic Inquiry* 39(4). 565–587.

Hunt, David. 2013. *Girt: The unauthorised history of Australia.* Vol. 1. Collingwood, Australia: Black.

Keller, Frank & Antonella Sorace. 2003. Gradient auxiliary selection and impersonal passivization in German: An experimental investigation. *Journal of Linguistics* 39(1). 57–108.

MacWhinney, Brian. 1992. Transfer and competition in second language learning. *Advances in psychology* 83. 371–390.

MacWhinney, Brian, Elizabeth Bates & Reinhold Kliegl. 1984. Cue validity and sentence interpretation in English, German and Italian. *Journal of Verbal Learning and Verbal Behavior* 23(2). 127–150.

Mohanan, Tara. 1994. *Argument structure in Hindi.* Stanford, CA: CSLI Publications.

Philipp, Markus, Matthias Schlesewsky & Ina Bornkessel-Schlesewsky. In preparation. To "be" or not to "be": ERP-correlates of auxiliary selection in German and the role of social cognition.

Scott, Graham. 1978. *The Fore language of Papua New Guinea.* Canberra, Australia: Pacific Linguistics.

Wang, L., Kumal Kumar Choudhary, Matthias Schlesewsky & Ina Bornkessel-Schlesewsky. In preparation. Actor-undergoer asymmetries in learning case-marking strategies.