used to rule out effects as a result of irrelevant properties of the experimental items (e.g. particular lexical items). It is our hope that strengthening methodological standards in the fields of syntax and semantics will bring these fields closer to related fields, such as cognitive science, cognitive neuroscience and computational linguistics.

References

1 Edelman, S. and Christianson, M. (2003) How seriously should we take Minimalist syntax? *Trends Cogn. Sci.* 7, 60–61
2 Schütze, C. (1996) *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*, University of Chicago Press
3 Cowart, W. (1997) *Experimental Syntax: Applying Objective Methods to Sentence Judgments*, Sage Publications
4 Wasow, T. and Arnold, J. (2005) Intuitions in linguistic argumentation. *Lingua* 115, 1481–1496
5 MacDonald, M.C. *et al.* (1994) The lexical nature of syntactic ambiguity resolution. *Psychol. Rev.* 101, 676–703
6 Gibson, E. and Pearlmutter, N. (1998) Constraints on sentence comprehension. *Trends Cogn. Sci.* 2, 262–268
7 Clark, H.H. (1973) The language-as-fixed-effect fallacy: a critique of language statistics in psychological research. *J. Verbal Learn. Verbal Behav.* 12, 335–359
8 Evans, J.S. *et al.* (1983) On the conflict between logic and belief in syllogistic reasoning. *Mem. Cogn.* 11, 295–306
9 Nickerson, R.S. (1998) Confirmation bias: a ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* 2, 175–220
10 Kayne, R. (1983) Connectedness. *Linguist. Inq.* 14, 223–249
11 Pesetsky, D. (2000) *Phrasal Movement and its Kin*, MIT Press
12 Clifton, C., Jr *et al.* (2006) Amnestying superiority violations: processing multiple questions. *Linguist. Inq.* 37, 51–68
13 Fedorenko, E. and Gibson, E. Adding a third *wh*-phrase does not increase the acceptability of object-initial multiple-*wh*-questions. *Syntax* (in press), doi:10.1111/j.1467-9612.2010.00138.x

doi:10.1016/j.tics.2010.03.005 Trends in Cognitive Sciences 14 (2010) 233–234

Letters Response

Quantitative methods alone are not enough: Response to Gibson and Fedorenko

Peter W. Culicover[1] and Ray Jackendoff[2]

[1] Department of Linguistics, The Ohio State University, 222 Oxley Hall, 1712 Neil Ave., Columbus, OH 43210-1298, USA
[2] Center for Cognitive Studies, Tufts University, Medford, MA 02155, USA

Gibson and Fedorenko [1] (see also [2,3]) correctly point out that subjective judgments of grammaticality are vulnerable to investigator bias, and that – where feasible – other types of data should be sought that shed light on a linguistic analysis. Major theoretical points often rest on assertions of delicate judgments that prove not to be uniform among speakers or that are biased by the writer's theoretical predispositions or overexposure to too many examples.

Another problem with grammaticality judgments is that linguists frequently do not construct enough control examples to sort out the factors involved in ambiguity or ungrammaticality. But this problem cannot be ameliorated by quantitative methods: experimental and corpus research can also suffer from lack of appropriate controls (see Box 1).

Nevertheless, theoreticians' subjective judgments are essential in formulating linguistic theories. It would cripple linguistic investigation if it were required that all judgments of ambiguity and grammaticality be subject to statistically rigorous experiments on naive subjects, especially when investigating languages whose speakers are hard to access. And corpus and experimental data are not inherently superior to subjective judgments.

In fact, subjective judgments are often sufficient for theory development. The great psychologist William James offered few experimental results [4]. Well-known visual demonstrations such as the Necker cube, the duck-rabbit, the Kanizsa triangle, Escher's anomalous drawings, and Julesz's random-dot stereograms are quick and dirty experiments that produce robust intuitions [5]. These phenomena do not occur in nature, so corpus searches shed no light on

Box 1. The need for proper controls in Gibson and Fedorenko's experiment

Fedorenko and Gibson's argument turns on the claim that superiority violations with two wh-phrases are supposedly worse than with three. Their experiment [9] disputes this judgment. The relevant sentence types are illustrated in (i).

(i) Peter was trying to remember ...
  a. who carried what.
  b. who carried what when.
  c. what who carried.
  d. what who carried when.

They find that, in contrast to longstanding judgments in the literature, (ic) is worse than (id), the two are judged to have about equal (un)acceptability.

They do not control by replacing the third wh-phrase with a full phrase as in (ii).

(ii) Peter was trying to remember ...
  a. who carried what last week.
  b. what who carried last week.

We find (iia) as good as (ia,b), but (iib) worse than (ic,d). If so, *some* violations with two wh-phrases *are* worse than counterparts with three. The difference calls for a reexamination of the examples in the literature, controlling for this factor. Ratings studies might be helpful in establishing the reliability of these judgments. We doubt relevant examples will be found in corpora of natural speech and writing. And we also doubt that Bolinger's original observation in [10] resulted from investigator bias.

Corresponding author: Jackendoff, R. (ray.jackendoff@tufts.edu).

them. Experiments can refine our understanding of such phenomena but it takes demonstrations to inspire us to design the experiments. Similarly, grammaticality judgments are the raw material for hypotheses about the structure of the language faculty. Without such judgments, the experimental enterprise cannot get off the ground.

Nor are experiments necessarily better than subjective judgments at distinguishing the factors affecting grammaticality. For example, if judgments reflect both grammatical structure and lexical class, failing to control for either will produce misleading results, regardless of how the judgments are arrived at [6]. Moreover, the question being asked can produce an uninformative answer, for example when [7] found that ranking produces gradient results for categorical concepts.

Corpora too can be useful, but it is not always clear what they tell us. Very frequent constructs in corpora are usually judged unproblematic by native speakers. Rare or variable constructs are the ones on which native speakers differ and theoretical issues often turn. Moreover, corpora can be contaminated by material from nonnative speakers. Finally, as [8] points out, the absence of a construct in a corpus can reflect many factors, including grammatical impossibility, inappropriateness of style, and processing complexity (think of multiple center-embedding). Again, it requires imagination to incorporate suitable controls.

We conclude that, as in all scientific inquiry, grammaticality judgments should be used as carefully as possible, controlling for all possible relevant factors (including confirmation bias), and that they should not be considered privileged over other sorts of data except by virtue of their convenience.

References

1 Gibson, E. and Fedorenko, E. (2010) Weak quantitative standards in linguistics research. *Trends Cogn. Sci.* 14, 233–234
2 Myachykov, A. *et al.* (2005) Attention and empirical studies of grammar. *The Linguist. Rev.* 22, 347–364
3 Ferreira, F. (2005) Psycholinguistics, formal grammars, and cognitive science. *The Linguistic Review* 22, 365–380
4 James, W. (1890/1950) *Principles of Psychology*, Dover Books
5 Jackendoff, R. (2007) Linguistics in Cognitive Science: The State of the Art. *The Linguistic Review* 24, 347–401
6 Schütze, C.T. (1996) *The Empirical Base of Linguistics*, University of Chicago Press
7 Armstrong, S. *et al.* (1983) What some concepts might not be. *Cognition* 13, 263–308
8 Yang, C. (2008) The great number crunch. *Journal of Linguistics* 44, 205–228
9 Fedorenko, E. and Gibson, E. Adding a third *wh*-phrase does not increase the acceptability of object-initial multiple-*wh*-questions. *Syntax.* (in press), doi:10.1111/j.1467-9612.2010.00138.x
10 Bolinger, D. (1978) Asking more than one thing at a time. In *Questions* (Hiz, H., ed.), pp. 107–150, Springer-Verlag

doi:10.1016/j.tics.2010.03.012 Trends in Cognitive Sciences 14 (2010) 234–235

Letters

Manipulating visual experience: Comment on Op de Beeck and Baker

Isabel Gauthier[1], Alan C-N. Wong[2] and Thomas J. Palmeri[1]

[1] Department of Psychology, Vanderbilt University, Wilson Hall, 111 21st ave South, Nashville, TN 37240, USA
[2] Department of Psychology, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong

In their recent *TiCS* contribution, Op de Beeck and Baker [1] (hereafter OB) suggest we abandon the idea that fairly local changes, limited to a single visual area, support visual learning. Instead, they propose that visual experience causes moderate and distributed changes that modulate pre-existing representations. We argue that their review overlooked something crucial: The *kind* of experience matters to how we learn visually. Unlike OB, we believe that *both* local and distributed changes can accompany visual object learning, depending on the task demands during learning.

OB review studies that use a wide variety of training tasks. For instance, participants (humans or monkeys) learn to categorize objects in one particular way [2], learn to discriminate visually similar objects [3] or learn to individuate objects by associating them with individual labels [4]. By focusing on the common aspects of visual learning, OB fail to note the potential importance of these training differences. This oversight is hard to avoid given current evidence. Indeed, many studies contrast categories that differ in shape, so we know that shape matters to the visual system [5]. Many studies hold shape constant but vary what participants are asked to attend to, so we know that attention can modulate visual responses [6]. And many studies, reviewed by OB, show that experience of some sort can change visual representations [2–4,7,8]. But because almost none of these studies *manipulate experience*, we have failed to learn much about whether the kind of experience with objects matters or not. As acknowledged by OB, most studies cannot even conclude if the learning effects obtained were a result of the complex training tasks or mere exposure to objects.

Recent work of ours manipulated experience by training different groups of participants with the same objects called ‘Ziggerins’ (Figure 1) for the same amount of time, but in very different ways [9,10]. One group learned to individuate Ziggerins by associating them

Corresponding author: Gauthier, I. (isabel.gauthier@vanderbilt.edu).