Language, abstraction and embodiment: A reply to Aronoff et al.

Polarization these days is rampant, and language researchers are not immune to the Zeitgeist. In linguistics, as in cognitive science, party-lines run along the "abstraction"-"embodiment" debate, a distant vestige of the bloody nature-nurture wars.

Mind-body, here, does not refer to Descartes' Dualism. Indeed, cognitive science assumes that *all* mental states are brain (i.e., bodily) states. So regardless of whether you endorse abstract rules or deny they exist, as a cognitive scientist, you are committed to the claim that *all* causes of behavior lie within the human body. Quibbles about whether cognition is "embodied", then, are a misnomer. What these debates are really about is whether some causes of cognition lie "below the head", so to speak: whether cognition is linked to *sensory* and *motor* bodily functions.

In its extreme form, the "embodiment" position reduces cognition to sensorimotor constraints; it asserts that there is no such thing as abstract concepts and representations[1, 2]. Your brain has no symbol for "cup" in general. Instead, your notion of a "cup" amounts to the sum of your sensations of specific cups and your motor interactions with them—how a cup feels as you hold it in your palms: its smoothness, roundness, coldness, and its color as registered by your eyes.

By the same token, your knowledge of language structure (e.g., why do you *blog*, not *lbog?*) is governed not by algebraic rules but by sensorimotor restrictions (e.g., *lbog* is banned because it is harder for your mouth to utter and for your ear to discern). And since our appreciation that "*lbog* is hard" only arises once our lips and tongues utter these sounds, it would seem that these "embodied" restrictions on language structure arise from experiences alone.

Whether language is in fact embodied has been the topic of active debate in linguistics and cognitive science[3-7]. But unfortunately, the notions of "embodiment" and "abstraction" are rarely spelled out. Aronoff and colleagues' passionate defense of "embodiment" does little to correct this problem—they never really tell us what "embodiment" really means. So when basic concepts are blurred, and the blood gets boiling, positions on this nuanced question can get needlessly polarized.

In the heat of battle, allegiances must be clear-cut. If you happen to conclude that speakers follow abstract rules of language, then you are automatically seen as asserting that these rules are entirely senseless and arbitrary; they can serve no functional purpose with respect to the transmission of language by the human body; as an extra bonus, some might falsely accuse you of stating that these rules must be innate. Conversely, if you believe that the design of the language system is adaptive, then, strangely enough, you are often seen as stating that language is the product of nurture, and that rules are utterly fruit of the imagination; language is entirely governed by the whims of "the body".

This is not a good conversation starter. So in the interest of opening up a more productive discussion, let me try to clarify these issues.

***

Aronoff et al. are right to ascribe to me the position that *some* rules of language, including some rules of phonology, are algebraic, amodal and abstract. They are wrong, however, to conclude that, from this, it follows that the design of language (especially, of phonology) is entirely arbitrary—that "embodied" constraints play no role—I assume that's what they mean by their assertion that "the body is intrinsic to the mental system of language, and not external to it".

How can one have it both ways, you wonder? The answer lies in three critical distinctions—scope (of one's argument), cognitive causation (proximal vs. distal), and grain size (of linguistic units). Let me spell them out.

I assert that *some* rules of language are amodal, inasmuch as they apply to both speech and sign. So when a naïve English speaker (for instance) who knows nothing about a sign language sees signs for the first time, they can apply some of the rules from their spoken language to extract the linguistic structure of those signs[8-10].

I do not assert that *all* linguistic rules can transfer across language modalities, and I certainly recognize that many rules of phonology differ across language modalities. A phonological restriction on voicing, for instance, has no business applying to manual signs—this much should be obvious. Claims to the contrary suffer from confusion about the *scope* of my assertion.

Nonetheless, I do assert that, at least in one case—that of doubling—the relevant rule is amodal. This discovery is significant, because it demonstrates that the computational machinery of language is algebraic[11, 12]. Just as the operation Y=2X references a class (e.g., X=any integer), as opposed to specific instances (e.g., *1,2,3…*), so does the restriction on doubling (e.g., "avoid XX", where X is any syllable) refers to a broad class ("any syllable") rather than to specific instances (*dada, baba* or  ,  , ). So as long as an English speaker can spontaneously recognize a signed syllable (which they demonstrably can[13]), and represent the formal function of identity, XX (which even newborns do[14]), the doubling rule from the English grammar ought to freely apply to ASL signs (as indeed it does [8-10]).

In sharing this finding with the readers of *Inference*, I sought to counter the pervasive popular belief that language (especially, phonology) is all about "talking": that one's linguistic intuitions (e.g., *lbog sounds "funny"*) are determined solely by the mechanics of our lips and tongue. This "folk phonology" runs so deep that people often take it for granted. The counter-intuitive phenomenon of cross-modal transfer debunks this myth.

This proposal also counters the claims of Aronoff and colleagues in two important respects. First, I do not vaguely claim that rules in the two modalities are somehow "similar" (as these authors suggest). Instead, I am committed to the strong claim that (in the case of doubling), the rule is precisely are *one and the same*. Second since algebraic variables

encode structure (XX), not speech instances (e.g., *baba*), the rule treats manual signs and aural speech alike. And this seems to fly directly in the face of Aronoff and colleagues' claim that "the body" shapes language structure.

But in reality, it may not. To explain why, we need to take a closer look at how "the body" can play a role in the cognition, generally, and the language system, specifically. Here, the distinction between "proximal" and "distal" causation is critical.

The algebraic hypothesis defines ***proximal*** causes of linguistic intuitions. In this view, people state that *laflaf* sounds "funny" because *lalaf* violates the *XX rule—the rule, then, is a proximal cause of their intuitions. Still, why do languages adopt such a rule in the first place?

Here, embodiment could matter ***distally***. Repetition indeed taxes the perceptual and motor systems [15, 16], so it stands to reason that repetition in language is going to be highly "regulated". These bodily constraints could provide the impetus—a distal cause that determine what rules "make it" into universal grammar in the course of language evolution. Distal bodily pressures, then, could "beget" universal grammar, which, in turn, could "beget" the doubling rule which drives speakers' intuitions. Crucially, once a rule "makes it" into one's grammar (e.g., of English), it is now the rule that is doing the "talking". As such, it is the rule, not bodily pressure, that is the proximal cause of speakers' intuitions. Still, bodily pressures matter distally[17, 18].

And of course, going beyond the grammar, the language system could also include non-grammatical analog mechanisms that are heavily embodied, and those can demonstrably shape speakers' intuitions. The phonetic system is a case in point.

Phonetics is a transducer. Spoken language phonetics (for instance) takes speech, which is analog and continuous, and outputs discrete categories, such as the voicing distinction between *bee* and *pee.* A large literature shows that this process is highly embodied: when your brain seeks to determine what you hear (*bee* or *pea*), it literally simulates (enacts) how you would *articulate* these speech sounds. We know this because, when people hear *bee*, they activate the motor area of the brain that controls the lips; when they hear *tea*, it's the area that controls the tongue that fires[19]. And when these areas are stimulated (by transcranial magnetic simulation, TMS), listeners' perception of these sounds changes accordingly[20, 21]. Work from my lab has shown that your perception of voicing in *pea* and *tea* changes even when you slightly press on your lips (relative to pressing on your tongue) [22]. All these demonstrations show that the articulatory motor system plays a direct causal role in the phonetic categorization of speech sounds. And here, "the body" is the direct proximal cause of perception, not merely a distal one.

But what's true within phonetics—outside the grammar—may not be true for the phonological rules within it. Here, you can "mess" up with the lips all you want (either directly, by pressing on them, or by stimulation the lip motor area, via TMS)—the application of phonological rules (concerning syllable structure) remains unaffected[23,

24]. This, to clarify, does not show that phonological rules are arbitrary; as noted, the body still plays a role. But this role of the body is apparently a distal one.

***

That language comes at different grain sizes (e.g., phonetics vs. phonology), which can differ with respect to their level of embodiment, should come at no surprise. Likewise, it should be only expected that, within the grammar (say, in phonology), some rules might be "incentivized" by distal bodily pressures. This makes perfect sense, given that phonology has a double duty to follow: it needs to generate novel forms by combinations, but it also needs to transmit them by relying on the human body. The solution is to favor rules that make bodily "sense"—this is exactly what would be expected of an adaptive system[17, 18].

This proposal allows for the possibility that some algebraic rules could apply amodally, whereas others could differ across language modalities, just as Aronoff et al. point out. Whether these modality differences could further shape the computations that are attested in speech and signs (as Aronoff et al suggest) is an interesting question—this remains to be seen.

How the language system arose in humans is another critical question that I will not consider here. Arnoff et al. submit that phonology cannot be innate, as it does not emerge fully *de novo*, in the first generation of emerging sign languages[25]. I'm not sure this follows. Innate systems need not be fully assembled at birth, immune to epigenetic triggers. Birdsong shows how a quintessentially innate system of communication could emerge gradually, across generations, through complex interactions of nature and nurture[26].

The nuanced view of embodiment (and innateness) I've painted here has some concrete methodological implications. When one seeks to evaluate the role of embodiment and abstraction in the language system, one ought to be proceed with caution. As noted, embodiment can play distinct roles at different level of analysis, and this role can be either proximal or distal. So when Aronoff and colleagues outline the undeniable correlations between the design of language and bodily pressures, it is irresponsible to jump to conclusions about causation—that the body is "internal" to the language system (presumably, as a proximal cause of language structure). As we know too well, correlations causations aren't one and the same. So rather than simply ask "is language embodied", the more appropriate question is "how". It's time to move beyond partly line and do the hard work of sorting this out.

1. Glenberg, A.M., J.K. Witt, and J. Metcalfe, *From the Revolution to Embodiment: 25 Years of Cognitive Psychology.* Perspectives On Psychological Science: A Journal Of The Association For Psychological Science, 2013. **8**(5): p. 573-585.
2. Barsalou, L.W., *Grounded cognition.* Annual Review of Psychology, 2008. **59**: p. 617-645.

3. Mahon, B.Z. and A. Caramazza, *A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content.* Journal Of Physiology, Paris, 2008. **102**(1-3): p. 59-70.
4. Barsalou, L.W., *Abstraction in Perceptual Symbol Systems.* Philosophical Transactions: Biological Sciences, 2003. **358**(1435): p. 1177-1187.
5. Leshinskaya, A. and A. Caramazza, *For a cognitive neuroscience of concepts: Moving beyond the grounding issue.* Psychonomic Bulletin & Review, 2016. **23**(4): p. 991-1001.
6. Bedny, M., A. Caramazza, E. Grossman, A. Pascual-Leone, and R. Saxe, *Concepts are more than percepts: The case of action verbs.* Journal of Neuroscience, 2008. **28**(44): p. 11347-11353.
7. Hauk, O., I. Johnsrude, and F. Pulvermüller, *Somatotopic Representation of Action Words in Human Motor and Premotor Cortex.* Neuron, 2004. **41**(2): p. 301-307.
8. Berent, I., O. Bat-El, D. Brentari, A. Dupuis, and V. Vaknin-Nusbaum, *The double identity of linguistic doubling.* Proceedings of the National Academy of Sciences, 2016. **113**(48): p. 13702-13707.
9. Berent, I., O. Bat-El, D. Brentari, and M. Platt, *Knowledge of language transfers from speech to sign: Evidence from doubling.* Cognitive Science, 2020(44): p. 1.
10. Berent, I., O. Bat-El, Q. Andan, D. Brentari, and V. Vaknin-Nusbaum, *Amodal phonology.* Journal of Linguistics, 2020: p. 1-30.
11. Chomsky, N. and M.P. Schützenberger, *The Algebraic Theory of Context-Free Languages*, in *Studies in Logic and the Foundations of Mathematics*, P. Braffort and D. Hirschberg, Editors. 1963, Elsevier. p. 118-161.
12. Marcus, G., *The algebraic mind: Integrating connectionism and cognitive science*. 2001, Cambridge: MIT press.
13. Berent, I., A. Dupuis, and D. Brentari, *Amodal aspects of linguistic design.* PLoS ONE, 2013. **8**(4).
14. Gervain, J., I. Berent, and J. Werker, *Binding at birth: Newborns detect identity relations and sequential position in speech.* Journal of Cognitive Neuroscience, 2012. **24**(3): p. 564-574.
15. Kanwisher, N., G., *Repetition blindness: type recognition without token individuation.* Cognition, 1987. **27**(2): p. 117-143.
16. Soto-Faraco, S. and C. Spence, *Modality-specific auditory and visual temporal processing deficits.* The Quarterly Journal Of Experimental Psychology. A, Human Experimental Psychology, 2002. **55**(1): p. 23-40.
17. Berent, I., *The phonological mind*. 2013, Cambridge: Cambridge University Press.
18. Berent, I., *The phonological mind.* Trends In Cognitive Sciences, 2013. **17**(7): p. 319-327.
19. Pulvermüller, F., M. Huss, F. Kherif, F. Moscoso del Prado Martin, O. Hauk, and Y. Shtyrov, *Motor cortex maps articulatory features of speech sounds.* Proceedings of the National Academy of Sciences, 2006. **103**(20): p. 7865-7870.
20. Möttonen, R. and K.E. Watkins, *Motor representations of articulators contribute to categorical perception of speech sounds.* The Journal of Neuroscience, 2009. **29**(31): p. 9819-9825.

21. Smalle, E.H.M., J. Rogers, and R. Möttönen, *Dissociating Contributions of the Motor Cortex to Speech Perception and Response Bias by Using Transcranial Magnetic Stimulation.* Cerebral Cortex, 2014.
22. Berent, I., M. Platt, R. Theodore, E. Balaban, P. Fried, and A. Pascual-Leone, *Speech perception triggers articulatory action:Evidence from mechanical stimulation.* Frontiers in communication, 2020.
23. Zhao, X. and I. Berent, *The basis of the syllable hierarchy: articulatory pressures or universal phonological constraints?* Journal of Psycholinguistic Research, 2018. **47**(1): p. 29-64.
24. Berent, I., A.-K. Brem, X. Zhao, E. Seligson, H. Pan, J. Epstein, A.M. Galaburda, and A. Pascual-Leone, *Role of the motor system in language knowledge.* Proceedings of the National Academy of Sciences, 2015. **112**: p. 1983-1988.
25. Sandler, W., M. Aronoff, I. Meir, and C. Padden, *The gradual emergence of phonological form in a new language.* Natural Language and Linguistic Theory, 2011. **29**: p. 505-543.
26. Fehér, O., H. Wang, S. Saar, P.P. Mitra, and O. Tchernichovski, *De novo establishment of wild-type song culture in the zebra finch.* Nature, 2009. **459**(7246): p. 564-568.