2025. Proc Ling Soc Amer 10(1). 5941. https://doi.org/10.3765/plsa.v10i1.5941.

The strength of conservativity: evidence from learnability experiments

Tyler Knowlton, John Trueswell & Anna Papafragou*

Abstract. One of the best-known semantic universals is determiner conservativity, roughly, the idea that the truth of sentences like *most fish swim* depends only on the things named by the determiner's first argument (*fish*). Conservativity has been argued to reflect a fundamental property of grammatical architecture, a conclusion bolstered by evidence that adults (and children) can pair novel determiners with conservative meanings but not with minimally-different non-conservative meanings. Here we extend this work to 'weakly' conservative meanings (ones for which the truth of the quantificational sentence depends either on the determiner's first argument or on its second argument). Such meanings are classically non-conservative but would be permitted under a weakened version of the generalization, which was designed to accommodate meanings like those of *only* and *even*. We find that adults cannot learn these kinds of novel determiner meanings either. This result suggests that the classical understanding of conservativity (on which 'weakly' conservative meanings are ruled out by virtue of being non-conservative) better describes the constraint that learners embody.

Keywords. conservativity; learnability; quantifiers; determiner meanings

1. Introduction. Determiner conservativity is a robust and renowned cross-linguistic generalization about the kinds of meanings natural language determiners can have (Barwise & Cooper 1981; Higginbotham & May 1981; Keenan & Stavi 1986). Informally, the generalization is that the truth of a quantificational sentence depends only on the things named by the determiner's first/internal/NP argument (i.e., the Noun Phrase with which the determiner initially combines). This is often abbreviated by saying that determiners 'live on' their first arguments.

For example, the determiners in (1) all 'live on' their first argument in the sense that only frogs and their properties matter for the truth of the sentence.

(1) All/Most/Some/No/... frogs are green.

One needn't look beyond the frogs to evaluate (1). Other green things – paint; trees; apples; aliens; etc. – are irrelevant. Put another way, conservative determiners can be thought of as restricting the domain to just the extension of their first argument (e.g., just the frogs, in this case).

In contrast, consider replacing one of the determiners in (1) with *only*, which is not a determiner, but often gets used prenominally in a way that makes it look like a determiner. In a sentence like *only frogs are green*, things that aren't frogs matter too. One can't figure out the truth of that sentence by restricting oneself to the frogs. Consequently, *only* doesn't 'live on' its first argument and thus isn't conservative. And while languages have words like *only*, the generalization is that they lack determiners with meanings like *only*'s. This bears repeating: If *only* were a

* For helpful discussion, we thank Jeffrey Lidz, Paul Pietroski, Alexander Williams, Zoe Ovans, and members of The Language and Cognition Lab at the University of Pennsylvania, The Language Learning Lab at the University of Pennsylvania, and The Perception and Language Lab at the University of Delaware. Authors: Tyler Knowlton, University of Delaware (tzk@udel.edu), John Trueswell, University of Pennsylvania (trueswel@psych.upenn.edu) & Anna Papafragou, University of Pennsylvania (anna4@sas.upenn.edu).

determiner (perhaps pronounced "donly"), then it would be a counterexample to determiner conservativity. But there are independent reasons for thinking *only* is not a determiner (including its syntactic promiscuity and focus-sensitivity; see e.g., Herburger 2000). And given that determiner conservativity is a generalization about determiner meanings (not meanings writ large), the fact that *only* is non-conservative does not threaten the robustness of the generalization.

One reason that conservativity has attracted so much attention is likely that it is so robust. Potential counterexamples have been proposed and explored, but the generalization seems to hold up in the face of such scrutiny (e.g., Herburger 1997; Romero 2018; Pasternak & Sauerland 2022). Indeed, determiner conservativity is perhaps one of the only genuine universals in semantics (von Fintel & Matthewson 2008). In part because of its unique position in this regard, it has been argued to stem from an architectural constraint on mental grammar. That is, conservativity seems not to be the result of general cognitive or computational principles, communicative pressures, or historical accident, but a constraint that reflects the way the language faculty is organized (see e.g., Pietroski 2004; Piattelli-Palmarini 2008; Romoli 2015; Knowlton et al. 2021, 2023; Lasersohn 2021; Ludlow & Živanović 2022).

The idea that conservativity reflects a deep fact about grammatical architecture – and thus is revealing about the organization of mental grammar – has received support from experimental work showing that conservativity is tied to learnability (Hunter & Lidz 2013; Knowlton et al. 2022; *cf.* Spenader & de Villiers 2019). The upshot of this work is that hypothetical determiner meanings that are non-conservative – ones that would run afoul of the generalization – aren't in learners' hypothesis space (a fact which, in turn, is explained by appealing to independently motivated proposals about the nature of the human language faculty).

This means that determiner conservativity not only occupies a privileged place among meaning-based typological generalizations, but also that it has implications for both the nature of the language faculty and for language acquisition. In short, it is an important generalization for theorists of various persuasions, and changes to its content are likely to have far-reaching ramifications.

1.1. A PROPOSED WEAKENING OF CONSERVATIVITY. Zuber & Keenan (2019) propose one such change. In particular, they propose 'weakening' conservativity from the constraint that determiners 'live on' their first argument to the weaker constraint that they 'live on' either their first or their second argument (but not both or neither). Any determiner that is classically conservative (i.e., one that 'lives on' its first argument) will *a fortiori* also be 'weakly' conservative. So this new constraint permits all existing determiners.

But 'weak' conservativity also permits meanings that classical conservativity rules out. One example is roughly the meaning of *only*, abstracting away from its focus-sensitive properties. We'll call this determiner-version *donly*, as in (2).

(2) Donly frogs are green.
≈ the green things are completely included in the frogs
≈ anything that's green is also a frog

As discussed above, *donly* is classically non-conservative: It's not true that just the frogs matter for the truth of (2), so *donly* clearly doesn't 'live on' its first argument. But *donly* is 'weakly' conservative, since it does 'live on' its second/external/predicative argument. Namely, it is true that just the green things matter for the truth of (2); one needn't look beyond the green things to

verify the sentence. So while a determiner like *donly* is ruled out on the classical understanding of the conservativity constraint, it would be allowed on the proposed 'weak' version.

The same can be said of determiners that roughly have the meaning – sans any focus-sensitive aspects – of *even* or *also*, which we might call *deven* and *dalso* and whose meanings we might gloss as in (3).

(3) Deven/Dalso frogs are green.
≈ *the green things include both frogs and non-frogs*
≈ *some of the green things are frogs and some are not*

This is also true of a determiner with a meaning similar to that of *both*, but inverted. That is, if *both frogs are green* means "every frog is green and there are two frogs", then imagine a hypothetical determiner *thob* that quantifies over and restricts the cardinality of the second argument instead of the first, as in (4).

(4) Thob frogs are green.
≈ *there are two green things and they're both frogs*
≈ *only frogs are green and there are two green things*

The hypothetical determiners in (2-4) all can be said to 'live on' their second/external/predicative arguments. Thus, they are classically non-conservative but 'weakly' conservative (i.e., ruled out on the classical understanding of the constraint, but permitted on the proposed weakening). This might be a desirable result if (a) such determiners are found to exist cross-linguistically, (b) one is unconvinced by the reasons for thinking expressions like *only*, *even*, and *also* are not determiners, or (c) one wants to expand the scope of the conservativity generalization beyond detereminers to a broader class of quantificational meanings.

It should also be noted that both versions of conservativity – classical and 'weak' – rule out many hypothetical determiner meanings (and thereby explain their cross-linguistic lack of existence). Classic examples of hypothetical non-conservative determiners like *equi* (meaning "equinumerious") and *ident* (expressing the identity relation) are ruled out by virtue of them 'living on' neither argument (e.g., to verify whether the frogs are equinumerous with the green things, one can't restrict oneself to just the frogs or to just the green things). Likewise, consider the hypothetical non-conservative determiner *allnon*, which combines universal quantification and negation such that *allnon frogs are green* means "the non-frogs are all green". Such a determiner is non-conservative on both versions of the constraint because it 'lives on' the complement of its first argument (meaning it fails to either 'live on' its first or its second argument).

In any case, Zuber & Keenan (2019)'s proposed weakening invites the following question: Which version of the conservativity constraint – classical or 'weak' – do learners possess? That is, which version of conservativity is a better description of the constraint on learners' hypothesis space? We address this question by extending our past work on determiner learnability in adults to include a 'weakly' conservative meaning.

2. Adult learnability experiment. The current learnability experiment pitting the two versions of conservativity against each other builds on the method reported in Knowlton et al. (2022). In this simple novel determiner learning task, adult participants are asked to learn the meaning of a new word, *gleeb*, by seeing various examples of *gleeb* used in a sentence like *gleeb of the circles*

are blue, accompanied by a simple picture of blue and orange circles and squares. Participants in this version of the task were randomly assigned to one of three conditions, which differed only in the meaning of *gleeb* that they were taught. These conditions are depicted in (5).

(5) Gleeb of the circles are blue.

a. ≈ *the circles outnumber the blue circles by 1* (conservative on both versions)
b. ≈ *the circles outnumber the blue things by 1* (non-conservative on both versions)
c. ≈ *the blue things outnumber the blue circles by 1* ('weakly' conservative)

In the Conservative condition in (5a), *gleeb* had a meaning that was conservative on both versions of the constraint (i.e., one for which *gleeb* can be said to 'live on' its first argument). In the Non-conservative condition in (5b), *gleeb* had a meaning that was non-conservative on both versions of the constraint (i.e., one for which *gleeb* 'lives on' neither argument). And in the Weak condition in (5c), *gleeb* had a meaning that was classically non-conservative but nonetheless 'weakly' conservative (i.e., one for which *gleeb* 'lives on' its second argument). The 'weakly' conservative meaning used here can be thought of as inverting the arguments of the conservative meaning in (5a), just as *donly* can be thought of as the 'inverted-arguments' version of *every*.[1]

Participants (n=90 monolingual English speakers recruited from prolific) were taught one of these three meanings for *gleeb* by viewing four different affirmative examples and four different negative examples, as shown in Figure 1. They saw each of the examples two times, for a total of 16 training trials. These trials were designed to rule out hypotheses corresponding to existing English determiner meanings (e.g., participants could not succeed at the task by learning that *gleeb* meant *most* or *some*).

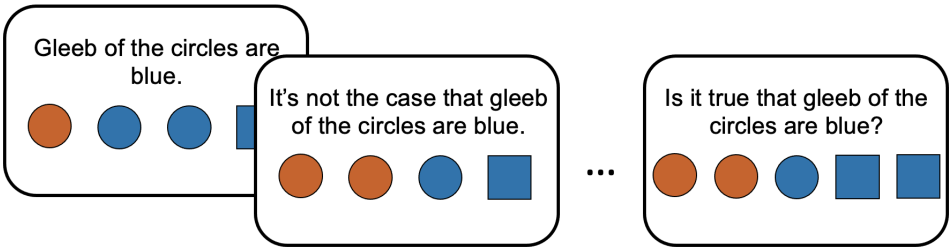

Figure 1. Example stimuli and experiment structure.

After 16 training trials, participants moved on to the test portion of the task, during which they were shown six new images (which were not used during training) and asked whether gleeb of the circles were blue with respect to those images. For three of these trials, the correct answer was "yes" and for three it was "no".

If the classical understanding of conservativity accurately describes the constraint that learners embody, then only participants in the Conservative condition should show evidence of learning (i.e., above chance performance on the test questions). We should expect the Non-conservative

[1] If *every frog is green* means "the frogs are completely included in the green things", then *donly frogs are green* can be glossed as "the green things are completely included in the frogs"; If *every frog is green* means "any thing that's a frog is green", then *donly frogs are green* can be glossed as "any thing that's green is a frog". In that sense, one could say that *every* and *donly* express the same relation, but with 'inverted' arguments.

and Weak conditions to pattern together, as both meanings are equally non-conservative from the point of view of classical conservativity. On the other hand, if 'weak' conservativity is a better description of the constraint on learners' hypothesis space, then participants should be equally able to learn in the Conservative and Weak conditions. Both views predict participants to struggle to learn in the Non-conservative condition (as the meaning in (5b) is ruled out by both versions of conservativity). The question, then, is whether performance in the Weak condition patterns with performance in the Non-conservative condition (as predicted by the classic understanding of conservativity) or with the Conservative condition (as predicted by Zuber & Keenan (2019)'s proposed weakening of conservativity).

2.1. RESULTS. As shown in Figure 2, only participants in the Conservative condition demonstrated any evidence of having learned the novel determiner. No evidence of learning was found in the Non-conservative or Weak conditions.

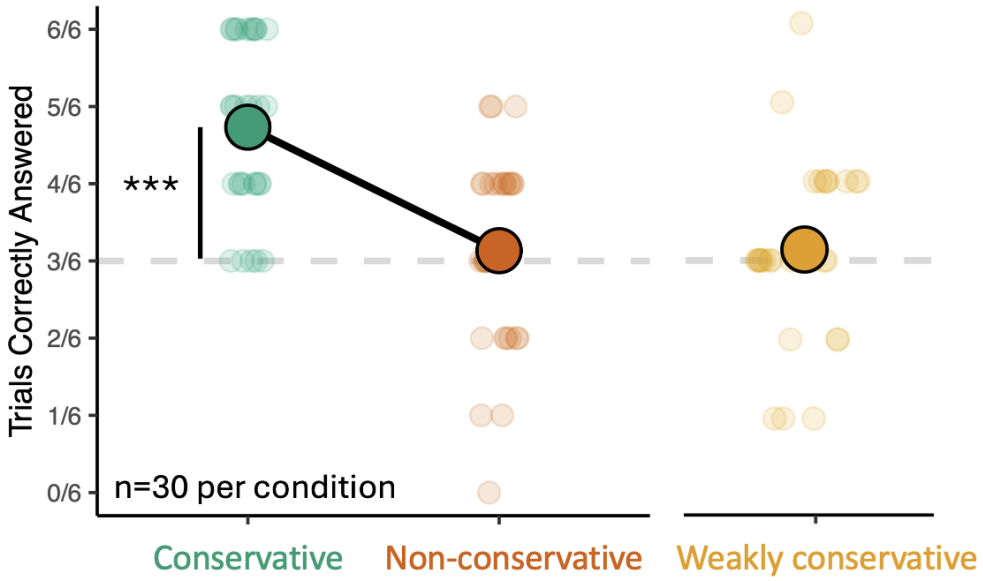

Figure 2. Average number of test trials correctly answered for each of the three conditions. Large points represent average performance, translucent points represent the number of test trials each individual participant correctly answered. Chance performance, 3/6, is represented by the grey dashed line. Stars represent significance of mixed-effects model intercepts and coefficients.

In particular, those in the Conservative condition performed above chance at test ($\beta = 1.72$ [95% CI: 1.22 to 2.22], $z = 3.43, p < .001$), whereas those in the Non-conservative condition did not ($\beta = 0.09$ [95% CI: -0.06 to 0.24], $z = 0.59, p = .553$), and neither did those in the Weak condition ($\beta = 0.09$ [95% CI: -0.08 to 0.26], $z = 0.55, p = .582$). Simply put, participants can't pair determiners with non-conservative meanings, and that includes meanings that are 'weakly' conservative.

3. Conclusion. Determiner conservativity is a semantic universal that is often argued to reflect a deep fact about the language faculty. It is an important generalization in part because it is potentially revealing about the nature of mental grammar. Here, we considered a recently proposed weakening of this generalization, which would make it slightly more permissive. However, the

results of our learnability experiment suggest that only meanings that conform to the stronger, classical understanding of conservativity are learnable; the 'weakly' conservative meaning we tested was apparently just as hard for participants to learn as an outright non-conservative meaning. This result suggests that the classical version of conservativity better describes the constraint that language learners bring to the table.

References

Barwise, Jon & Robin Cooper. 1981. Generalized quantifiers and natural language. In *Philosophy, Language, and Artificial Intelligence*, 241–301. Springer. https://www.jstor.org/stable/25001052.

Herburger, Elena. 1997. Focus and weak noun phrases. *Natural Language Semantics* 5(53). 53–78. https://doi.org/10.1023/A:1008222204053.

Herburger, Elena. 2000. *What counts: Focus and quantification*. MIT Press. https://doi.org/10.7551/mitpress/7201.001.0001.

Higginbotham, James & Robert May. 1981. Questions, quantifiers and crossing. *The Linguistic Review* 1(1). 41–80. https://doi.org/10.1515/tlir.1981.1.1.41.

Hunter, Tim & Jeffrey Lidz. 2013. Conservativity and learnability of determiners. *Journal of Semantics* 30(3). 315–334. https://doi.org/10.1093/jos/ffs014.

Keenan, Edward L & Jonathan Stavi. 1986. A semantic characterization of natural language determiners. *Linguistics and Philosophy* 9(3). 253–326. https://www.jstor.org/stable/25001246.

Knowlton, Tyler, Paul Pietroski, Alexander Williams, Justin Halberda & Jeffrey Lidz. 2021. Determiners are "conservative" because their meanings are not relations: Evidence from verification. *Semantics and Linguistic Theory* 30. 206–226. https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/30.206.

Knowlton, Tyler, Paul Pietroski, Alexander Williams, Justin Halberda & Jeffrey Lidz. 2023. Psycholinguistic evidence for restricted quantification. *Natural Language Semantics* 31(2). 219–251. https://doi.org/10.1007/s11050-023-09209-w.

Knowlton, Tyler, John Trueswell & Anna Papafragou. 2022. New evidence for the unlearnability of non-conservative quanti ers. In *Proceedings of the amsterdam colloquium*, 374–381.

Lasersohn, Peter. 2021. Common nouns as modally non-rigid restricted variables. *Linguistics and Philosophy* 44(2). 363–424. https://doi.org/10.1007/s10988-019-09293-4.

Ludlow, Peter & Sašo Živanović. 2022. *Language, form, and logic: In pursuit of natural logic's holy grail*. Oxford University Press. https://doi.org/10.1093/oso/9780199591534.001.0001.

Pasternak, Robert & Uli Sauerland. 2022. German measurement structures: Case-marking and non-conservativity. *The Journal of Comparative Germanic Linguistics* 25(2). 221–272. https://doi.org/10.1007/s10828-022-09134-y.

Piattelli-Palmarini, Massimo. 2008. Events and conservativity: Clues towards language evolution. *Knowledge, language, and interpretation: On the philosophy of Donald Davidson* 14. 81. https://doi.org/10.1515/9783110322521.81.

Pietroski, Paul. 2004. *Events and semantic architecture*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199244300.001.0001.

Romero, Maribel. 2018. The conservativity of *many*: Split scope and *most*. *Topoi* 37. 393–404. https://doi.org/10.1007/s11245-017-9477-5.

Romoli, Jacopo. 2015. A structural account of conservativity. *Semantics-Syntax Interface* 2(1). 28–57. https://pure.ulster.ac.uk/en/publications/a-structural-account-of-conservativity-3.

Spenader, Jennifer & Jill de Villiers. 2019. Are conservative quantifiers easier to learn? evidence from novel quantifier experiments. In *22nd Amsterdam Colloquium*, http://hdl.handle.net/11370/79051821-bdba-4168-8214-5cf1033b8451.

von Fintel, Kai & Lisa Matthewson. 2008. Universals in semantics. *The Linguistic Review* 25(1-2). 139–201. https://doi.org/10.1515/TLIR.2008.004.

Zuber, Richard & Edward L Keenan. 2019. A note on conservativity. *Journal of Semantics* 36(4). 573–582. https://doi.org/10.1093/jos/ffz007.