Variable Names Don't Matter (and aren't needed)

Pauline Jacobson
Dept. of Cogntive, Linguistic and Psychological Sciences
Brown University
pauline_jacobson@brown.edu

Meta-abstract

The literature contains several arguments for the existence of variable names (and hence variables) that assume a (arguably stipulative) principle of No Meaningless Coindexation (NMC) (e.g., Heim, 1997). For example, Takahashi and Fox (2005) (T&F) use variable names (supplemented with NMC) combined with 'MaxElide' to account for ellipsis facts first observed in Sag (1976). MaxElide (from Merchant, 2001) roughly blocks ellipsis of smaller material in certain cases if a larger ellipsis is possible. But this is not always the case - and T&F (following Merchant 2001) suggest that this holds only in cases involving unbound variables within a domain relevant to the conditions on ellipsis. I argue that (contra the claims in T&F) all of the facts are readily accounted for in a variable-free semantics. Here this is recast as a competition not between ellipsis *size* but *types*: <e,<e,t>> ellipsis is blocked when there is a competing ellipsis in which the missing material is of type <e,t>. More generally, ellipsis is blocked if there is a competing ellipsis of a simpler type. Type competition not only accounts for the observations in T&F, but also for the original Merchant 'MaxElide' observations. Moreover, type competition has a natural explanation in terms of a pressure to ease the processing burden of a listener, while T&F's competition principle has no obvious motivation. I will root this in the more general landscape of arguments for variable names using NMC vs. variable-free alternatives, showing that the latter have theoretical and empirical advantages.

Variable Names Don't Matter (and aren't needed)

Do Variable Names Matter? This paper revisits the question of whether variable *names* (and hence, of course, variables) play a role in the account of any grammatical phenomena (a question with a long history in *SALT* proceedings). The focus here is on cases that appear to support variables via a requirement that two different pronouns (or, traces) 'bound' by different binders must be different semantic objects. To illustrate schematically consider (1):

(1) Every third grade boy_i loves his_i mother. Every fourth grade boy_j hates his_j mother.

The claim at issue is that (in other cases of this general sort) the two occurrences of *his* must correspond to different semantic objects. Note that the existence of variable names ensures that the two occurrences of *his* can be different objects but this is not enough. Enter the principle of No Meaningless Coindexation (first formulated in Heim 1997 and given informally in (2), see also Takahashi and Fox 2005, Kennedy 2014) which ensures that they must be different objects.

(2) Two pronouns bound by different binders cannot be coindexed.

(2) guarantees that the two pronouns *his* in (1) have different indices, and so do not have the same value on all assignments. One way to enforce (2) is via a constraint that every 'binder' (in a single discourse) has a distinct index (Kennedy, 2014). Call the claim that the two pronouns in (1) must have different values the Names Matter hypothesis.

Compare this to a variable-free semantics (VFS) as advocated in, e.g., Jacobson (1999). Here both occurrence of *his* denote the identity function on individuals. Since there are no assignments, there is no way to require these to be different semantic objects.

The initial focus of this paper is on an argument for the Names Matter hypothesis (and hence for the existence of variables) in Takahashi and Fox (2005) (hereafter, T&F). T&F discuss an interesting set of facts surrounding ellipsis and give an account making crucial use of (2) (and hence of variable names). We show that these facts are readily accounted for in a variable-free semantics in a way that is at least as simple and natural. We then turn to a more general overview of the issue, arguing that there is no real evidence for names mattering and that recasting certain generalizations without use of variable names actually has benefits.

'MaxElide' and Variable Names. By way of background, Sag (1976) pointed out that cases like (3b) are degraded on the sloppy reading. (The strikethroughs are used to indicate the relevant reading, without commitment to the claim that there is actual silent linguistic materia.)

(3) a. $Cruz_i$ thinks that Iowa will vote for him_i, and $Rubio_j$ also does ~~think that Iowa will vote for him_j~~

b. ?*$Cruz_i$ thinks that Iowa will vote for him_i and $Rubio_j$ also thinks it will ~~vote for him_j~~.

The notation ?* is somewhat arbitrary as many speakers find improvement with the right context; see Grant (2008). Nonetheless, the contrast between (3a) and (3b) is real. We will not pursue Sag's account of this here, but will note that the degradedness of (3b) is surprising under the VFS account of pronouns and ellipsis. For example, Jacobson (1992) argues that Antecedent Contained Deletion cases are simply a matter of 'elided' or 'missing' meanings of type <e,<e,t>>. She referred to this as Transitive Verb Phrase Ellipsis or TVPE - and we use the term TVPE here to mean any ellipsis where the missing meaning is of type <e,<e,t>>. If 'ellipsis' can involve missing meanings of type <e,<e,t>> as well as of type <e,t>, then (3b) should be fine. The 'antecedent' is *vote for him* and the understood meaning (or, the meaning of silent material) in the second clause is the same - both are just the function λx[vote-for(x)] (i.e., [[vote-for]]). While there are many different view on VPE (and TVPE), almost all predict that (3b) this should

be good if the understood meaning in the Rubio clause is the same as the meaning of the 'antecedent' material in the Cruz clause.

And indeed, beginning with Evans (1988) it has been shown that there are plenty of good cases of this general sort. For example, (4) is much better:

(4) $Cruz_i$ thinks that Iowa will be the most likely to vote for him_i, while $CHRISTIE_j$ thinks that NEW HAMPSHIRE will ~~be the most likely to vote for him_j~~.

(Such examples may not be as robustly good as ordinary VPE; I return to this below. But the literature contains considerable agreement that examples of this type are possible,) Thus cases like (4) are to be expected under TVPE analysis, the 'antecedent' and the understood meaning in both the Cruz and Christie clauses is the 2-place relation λx[be the most likely to vote for x].

But this leaves unexplained the Sag contrasts. Various authors (e.g.T&F 2005, Merchant 2008) have made an important observation: (3b) competes with (3a). Thus T&F and Merchant derive the contrast from a MaxElide principle (cf., Merchant, 2001, 2008). Informally (and more generally than usually stated) this says that when both a 'small' and a bigger ellipsis are possible, there is a pressure to use the bigger ellipsis. Note that I have phrased this in terms of a pressure - presumably on a speaker - rather than as a transderivational grammatical principle (but this choice has no consequences for T&Fs analysis). Hence, (3b) is bad (or funny) because (3a) is available. But in (4) there is no possible bigger ellipsis because the subjects of the intermediate clause (Iowa and NH) are different. Hence no competition, and so (4) is better.

If this were the end of the story, this would provide no evidence for variable names: the claim that MaxElide (as given so far) is responsible for the contrasts is neutral between a theory with variables and the TVPE (variable-free) analysis. Thus the next piece crucial to the Names Matter account is a claim that the pressure for higher ellipsis exists *only* when there is an unbound pronoun (or trace) in the ellipsis site. (Merchant 2008) Consider (5):

(5) a. Denise thinks that Bernie can win, and Elizabeth also does ~~think B can win~~.

b. Denise thinks that Bernie can win, and Elizabeth also thinks he can ~~win~~.

(Actually one might think that there is an unbound *he* here in the larger ellipsis in (5a) since we find *he* in the fuller version (5b). The full paper will show (5a) is consistent with my final analysis.) T&F claim that (5b) is just fine. While all of these judgments are delicate, I agree that (5a) and (5b) show nothing like the Sag contrasts in (3). (If they do, then this is the end of the story: variable names are irrelevant here.) Hence T&F define the notion of a Parallelism Domain (PD) - which is tied in with Rooth's focus condition on VPE. Space precludes a full exposition; informally take a PD to be the *lowest* domain (containing an ellipsis site) for which there is an antecedent for that ellipsis with the same meaning *or* one which meets Rooth's focus condition. MaxElide consults only that PD. So (5b) is good because *win* is its own PD, and MaxElide will look no higher. But in (3b), *vote for* him_j is not its own PD as the attempted 'antecedent' is *vote for* him_i (note the crucial role of No Meaningless Coindexation here). The PD is thus the higher one going up to the Rubio clause (where the Cruz clause satisfies Rooth's focus condition). As they point out, a similar move is unavailable in the variable free TVPE account, because Rubio's [[vote for him]] is identical in meaning to Cruz's [[vote for him]].

Rethinking the competition. I will claim that the degradation of (3b) is indeed due to competition with (3a), and (4) is better because there is no competing case. But there is a simple alternative to stating the competition in terms of ellipsis size. Suppose that instead the competition is between an 'ellipsis' (or, a 'missing' meaning) or type <e,<e,t>> (the "TVPE ellipsis cases) and a missing meaning of type <e,t>. (Thus as in most accounts of sloppy readings in general beginning with Keenan 1971, we assume that the 'missing' or 'elided' material

in (3b) is the <e,t> function λx[x thinks Iowa vote for x].) More generally, suppose there is a pressure to elide simpler meaning types when available over complex ones, where an n-place function is simpler than an n+m place function. This is especially natural when coupled with a view that VPE (and TVPE) involves an auxiliary behaving like an anaphor (see Hardt, 1993, Jacobson 2008), and where these ellipsis phenomena simply involve a listener having to pick up a contextually salient meaning. (That meaning very often is made salient by having been named - hence the *illusion* of an antecedent.) But even if the *grammar* requires actual silent linguistic material in the ellipsis site, the fact is that the *listener* does not hear this material (in this position) and so must supply the meaning. (Thus, unless it turns out that processing involves the listener copying in an actual string - which has already been interpreted - and then reinterpreting it, it makes sense to view the pressure as one which eases the processing burden.) Assume further that more complex meanings are harder to access. Then the competition is a consequence of a speaker making things easier for the listener. That <e,<e,t>> meanings are more difficult to access than <e,t> meanings is consistent with the basic fact that cases like (4) are perhaps not perfect, and with the fact that it is very difficult to pragmatically supply <e,<e,t>> meanings which have not been named (but see Jacobson 2008 for such cases). Note that the fact that the judgments can be pushed around with context is consistent with this view, and we also will compare the consequences of rethinking this as type competition rather than size competition.

The above facts all follow. (3b) is funny due to competition with (3a), for (3a) involves a missing <e,t> meaning while (3b) involves a missing <e,<e,t>> meaning. (4) is better because there is no competitor. (5b) is fine because it does not compete with (5a) - both involve missing <e,t> meanings. Moreover, T&F note that the 'little' ellipsis is alright when (in their terms) it contains an unbound pronoun within it just in case that pronoun is cobound with the corresponding pronoun in the antecedent clause (in that case, the two have the same indices). But these cases also follows from the account here. Since the 'binding' is from higher up, both 'big' and 'little' ellipsis involve missing meanings of type <e,<e,t>> and hence there is no competition. (The full paper will examine some more intricate cases where the two accounts converge.)

There is a further pleasing consequence to this analysis: it extends directly to the contrasts which originally motivated the notion of MaxElide:

(7) Elizabeth will endorse one of those candidates, but I don't know which/*which she will.

In the view here, the material that is missing is in the VP ellipsis case of type <e,<e,t>> (it is a VP meaning with a 'gap') whereas the Sluicing is missing only a meaning of type <e,t>.

More generally. I will then locate this in a more general 'state of the art' look at the Names Matter hypothesis as it pertains to TVPE (aka 'rebinding') cases. For example, Heim (1997) used (2) to account for some puzzling facts due originally to Kennedy (1994), and Kennedy (2014) recently revives (and revises) Heim's analysis (Kennedy's revision is designed to solve some problems with Heim's original analysis). But Jacobson (2009) showed that Heim's basic insight (that Kennedy's 'puzzle' has to do with focus considerations) can be implemented without variable names whereby the insight is stated in purely semantic (variable-free) terms. Here we will show that her variable-free version actually has an interesting payoff: it accounts for the following asymmetry in TVPE noted in Merchant (2001) (under the rubric 'rebinding')

(8) a. John interviewed all of the prisoners. But only the lifers wanted him to.

b. John interviewed all of the lifers. ?But all of the prisoners wished he had.

(Context can improve (8b), which follows from the analysis.) The purely variable-free analysis of the interaction with focus predicts the contrast; it is not clear how variable names can help. More generally, then, names do not seem to matter (which follows if they do not exist!).

References

Evans, F. 1988. 'Binding into Anaphoric Verb Phrases' in *Proceedings of ESCOL* 5.

Grant, M. 2008. "A Psycholinguistic Investigation of MaxElide in variable binding contexts'. Abstract available at http://conf.ling.cornell.edu/nels39/NELS-39Abstracts/grant.pdf.

Hardt, D. 1993. *Verb Phrase Ellipsis: Form, Meaning and Processing*. Ph.D. Dissertation, University of Pennsylvania.

Heim, I. 1997. 'Predicates or Formulas? Evidence from Ellipsis', *Proceedings of SALT 7*.

Jacobson, P. 1992. 'Antecedent Contained Deletion in a variable-free semantics', *Proceedings of SALT 2*.

Jacobson, P. 1999. "Toward a Variable Free Semantics", *Linguistics and Philosophy* 22.

Jacobson, P. 2008. "Direct Compositionality and Variable Free Semantics: The Case of Antecedent Contained Deletion" in K. Johnson (ed.) *Topics in Ellipsis,* Cambridge University Press.

Jacobson, P. 2009. 'Do Representations Matter or Do Meanings Matter?: The Case of Antecedent Containment', in E. Hinrichs and J. Nerbonne (eds.), *Theory and Evidence in Semantics: Papers in Honor of David R. Dowty*. CSLI Publications.

Keenan, E. 1971. 'Names, quantifiers, and the sloppy identity problem', *Papers in Linguistics* 4.

Kennedy, C. 1994. 'Argument Contained Ellipsis'. ms., UC Santa Cruz. Published in K. Johnson (ed.), *Topics in Ellipsis,* Cambridge University Press (2008).

Kennedy, C. 2014. 'Predicates AND Formulas: Evidence from Ellipsis', presented at *SALT 14,* published in L. Crnic and U. Sauerland (eds), *The Art and Craft of Semantics: A Festschrift for Irene Heim,* MITWPL 70.

Merchant, J. 2001. *The Syntax of Silence*. Oxford University Press.

Merchant, J. 2008. 'Variable Island Repair Under Ellipsis', in K. Johnson (ed.), *Topics in Ellipsis,* Cambridge University Press.

Sag, I. 1976. *Deletion and Logical Form*. Ph.D. Dissertation, MIT.

Takahashi, S. and D. Fox 2005. 'MaxElide and the Rebinding Problem'. *Proceedings of SALT 15*.