

A probabilistic pragmatics for English singular *some*

Overview When English *some* takes a singular NP, the resulting phrases reliably convey meanings not shared by variants with *a(n)* or plural NPs. Thus, whereas *I met a friend* and *I met some friends* are unmarked, *I met some friend* is associated with a range of affective meanings, especially speaker lack of interest in the referent. Farkas (2002) and Weir (2012) trace these effects to semantic properties of *some* that entail that the speaker cannot identify a unique referent for the phrase (see also Alonso-Ovalle and Menendez-Benito 2003; Solt and Stevens 2018). Choosing *some* over its competitors is thus a choice to convey this uncertainty directly, which often leads to negative pragmatic inferences.

We identify two challenges for semantic accounts of singular *some*. First, we present attested examples in which speakers use singular *some* to refer to entities that they themselves identify in the context. Second, we argue that semantic accounts miss an important generalization: *some* conveys these special meanings only if *a* is available as competitor. These challenges suggest a pragmatic source for the relevant meanings. We propose to derive them as manner implicatures, using similar logic to that of Weir (2012) but without encoding a fixed semantic contrast between *a* and *some*. To make this precise, we adopt the version of the Rational Speech Acts (RSA) model developed by Bergen et al. (2016), in which pragmatic reasoning leads synonymous forms that differ in markedness to take on different usage patterns.

Challenges to semantic accounts For Farkas (2002), singular *some* imposes a discourse-semantic requirement that its associated discourse referent not be *identifiable in context*: the discourse referent should be compatible with multiple real-world entities. This notion of identifiability can be refined to hold separately for the speaker and listener(s); Farkas emphasizes the speaker-relative version. Weir's (2012) proposal is similar: singular *some* has an anti-singleton presupposition that helps derive the *differentiation condition*, which says that a speaker who uses singular *some* cannot “if presented with the extension of NP, ‘pick out’ the witness of the existential claim”. This is also a consequence of Farkas's discourse semantics.

In support of these speaker non-identifiability constraints, Farkas and Weir offer examples like the following, which we give with their original judgments:

- (1) Susan rented some movie for us yesterday. #It was *The Maltese Falcon*. (Farkas 2002: 70)
- (2) You are lost. You know that the city you're in has only two squares. You keep coming across both squares. You can tell them apart because one has a fountain and the other doesn't. You end up in the fountainless square in the city. Your friend phones you:
A: Where are you? B: I'm in a/#some square in the city. (Weir 2012: 182)

While we are unsure of the precise status of these specific examples, we agree that speaker non-identifiability seems robust across a range of uses. However, we find many attested examples that are similar in form to the above. The following are from CoCA (Davies 2008):

- (3) What does it mean for Kanye [...] to appear in a photograph in a MAGA hat with some guy – namely music industry executive Lyor Cohen (more about him later) [...]
- (4) In Julian Assange's world we are the bad guys [...] I hope the president-elect will get his information and trust the American patriots who work in the intelligence community who swear an oath and allegiance to the constitution and not some guy hiding from the law with a record of undercutting and undermining American democracy.
- (5) I saw some statue in Italy, and the expression on the guy's face blew me away. Like, I could read and recognise an emotion because some one 400 years ago felt how I do [...]

In (3)–(5), the referent is identified directly or clearly known to the speaker. Thus, speaker non-identifiability seems not to be a correct characterization. Nonetheless, the examples are pragmatically rich in a way that their variants with *a(n)* are not; *some* seems to convey irrelevance in (3), derision in (4), and generality in (5). The fluid nature of these inferences suggests that a purely pragmatic account is called for.

The importance of alternatives Even if we find a semantic denotation that is sufficient to account for examples like the above, we will have to specify its distribution very carefully. We have already noted that plural forms do not give rise to the effects of interest. The effects are also absent with mass nouns. And forms like *someone* and *something* seem to act like regular indefinites as well.

The fact that the semantics must distinguish singular *some* from the form or forms that participate in these other constructions is not unpalatable on its own. However, such an account misses a clear generalization: the effects for *some* arise only where *a(n)* is available. Since *a(n)* cannot combine with plurals or mass nouns, plurals are bleached of the semantic effects. Since *someone* and *something* are distributionally different from *a one/person* and *a thing*, the relevant competition is not in play there either. These facts too argue for a pragmatic account of singular *some*.

An RSA account In the version of RSA developed by Bergen et al. (2016), two synonymous expressions can convey different things if one is more marked. In such situations, the marked form associates with the less likely meaning, and the unmarked form associates with the more likely meaning (Horn 1984; Levinson 2000). The primary mechanism by which this happens is *semantic uncertainty*: speakers and listeners are modeled as reasoning not just about the world but also about the specific denotations that words and phrases should have in context. In this model, synonymous terms come apart even if they have exactly the same set of possible meanings, in virtue of their markedness contrasts. This model has been applied to a wide range of well-known pragmatic effects, including not only markedness implicatures but also I-implicatures (Potts and Levy 2015), embedded implicatures (Potts et al. 2016), and free-choice disjunction (Champollion et al. 2019).

To apply this model, we make four crucial assumptions. First, *some* is more marked than *a*. We support this indirectly using corpus statistics. For instance, in CoCA, singular *a* is 54 times more frequent than singular *some* (4,751,798 tokens to 87,290 tokens). Second, noun phrases can be used to signal presumptions about both speaker identifiability and listener identifiability. Third, *some* and *a* can in principle be used to signal any of these effects – in other words, they are synonymous and have all the same pragmatic refinements in the model. Fourth, states of speaker non-identifiability (\neg SI) are less likely than those of speaker identifiability (SI).

With these assumptions in place, the model delivers the results given in Figures 1a–1b. The central findings: for singular nouns, a speaker who wishes to signal \neg SI strongly prefers *some*, and a listener who hears *some* is strongly biased in favor of \neg SI states. In turn, regular indefinites take on a bias for mere listener non-identifiability (\neg LI), assuming that a proper name is a salient alternative. The dynamic is very different where *a* is not an alternative, as seen in Figures 1c–1d; here, even with all the same assumptions about markedness and likelihood in place, *some* is preferred by the speaker for the entire range of indefinite states, and listener inferences are simply proportional to the relative likelihood of those states as given by the prior.

Consequences The RSA account captures the core observations for singular *some*, characterizing them as a manner implicature arising from competition among synonymous forms with differing markedness. It also makes sense of the range of judgments seen in (1)–(5). In particular, (1)–(2) are certainly surprising things for a speaker to produce because of the bias against speaker identifiability that *some* will instill in the listener. However, there is no sense in which these discourses are incoherent. Indeed, (3)–(5) show that speakers can leverage this unexpectedness to achieve complex pragmatic and social effects ranging from lack of interest in the referent to outright disdain.

	a NP _{sg}	<i>some</i> NP _{sg}	PN
SI & LI	0.0	0.0	1.0
SI & ¬LI	0.94	0.06	0.0
¬SI & LI	0.16	0.84	0.0
¬SI & ¬LI	0.16	0.84	0.0

(a) With singular NPs, the pragmatic speaker chooses *some* with high probability given a ¬SI state, and *a* emerges as the preferred choice for conveying listener non-identifiability.

	<i>some</i> NP _{sg}	PN
SI & LI	0.0	1.0
SI & ¬LI	1.0	0.0
¬SI & LI	1.0	0.0
¬SI & ¬LI	1.0	0.0

(c) With plural NPs, the pragmatic speaker chooses *some* wherever the speaker or listener can't identify the referent; the special effects of *some* do not arise in the absence of competition with *a*.

	SI & LI	SI & ¬LI	¬SI & LI	¬SI & ¬LI
a NP _{sg}	0.0	0.92	0.04	0.04
<i>some</i> NP _{sg}	0.0	0.13	0.43	0.43
PN	1.0	0.0	0.0	0.0

(b) With singular NPs, pragmatic listener assumes that the state is a ¬SI one with high probability upon hearing *some*.

	SI & LI	SI & ¬LI	¬SI & LI	¬SI & ¬LI
<i>some</i> NP _{pl}	0.0	0.67	0.17	0.17
PN	1.0	0.0	0.0	0.0

(d) With plural NPs, the pragmatic listener cannot distinguish among the various indefinite states and so favors the more likely one given the state prior, which disfavors ¬SI by assumption.

Figure 1: RSA predictions for singular and plural NPs. All the predictions are probabilistic and meant to reflect relative preferences in production and interpretation. We use ¬SI & ¬LI to model the equivalence classes of states in which the relevant discourse referent is neither speaker nor listener identifiable, and similarly for the other state classes. PN is an appropriate proper name, which we specify as requiring both speaker and listener identifiability. We assume (i) the cost of *some* is greater than the cost of *a* (in the above, we make the cost of *some* twice that of *a*); (ii) the ¬SI states are less likely than the SI & ¬LI state (here, we illustrate with $P(\text{SI}\&\text{LI}) = P(\text{SI}\&\neg\text{LI}) = 0.4$ and $P(\neg\text{SI}\&\text{LI}) = P(\neg\text{SI}\&\neg\text{LI}) = 0.1$); and (iii) *some* and *a* phrases can denote any non-empty subset of the available states.

References

- Alonso-Ovalle, Luis and Paula Menendez-Benito. 2003. Some epistemic indefinites. In *NELS 33*, 1–12.
- Bergen, Leon; Roger Levy; and Noah D. Goodman. 2016. Pragmatic reasoning through semantic inference. *Semantics and Pragmatics* 9(20).
- Champollion, Lucas; Anna Alsop; and Ioana Grosu. 2019. Free choice disjunction as a rational speech act. In *SALT 29*.
- Davies, Mark. 2008. *The Corpus of Contemporary American English*. <http://corpus.byu.edu/coca/>.
- Farkas, Donka F. 2002. Varieties of indefinites. In *SALT 12*, 59–83.
- Horn, Laurence R. 1984. Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In *Meaning, Form, and Use in Context: Linguistic Applications*, 11–42.
- Levinson, Stephen C. 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. MIT Press.
- Potts, Christopher; Daniel Lassiter; Roger Levy; and Michael C. Frank. 2016. Embedded implicatures as pragmatic inferences under compositional lexical uncertainty. *Journal of Semantics* 33(4):755–802.
- Potts, Christopher and Roger Levy. 2015. Negotiating lexical uncertainty and speaker expertise with disjunction. In *BLS 41*.
- Solt, Stephanie and Jon Stevens. 2018. *Some Three Students: Towards a unified account of some*. In *SALT 28*, 345–365.

Weir, Andrew. 2012. *Some*, speaker knowledge, and subkinds. In *Proceedings of the ESSLLI 2012 Student Session*, 180–190.