

Semantic change

Lessons from theories of grammar change

Rob Truswell

EGG, Brno, August 2015

Introduction

- ▶ We can get part of a theory of semantic change for free by building on the fact that (compositional) semantics is part of grammar.
- ▶ So change in the compositional semantic system is a type of grammatical change.
- ▶ And there are theories of grammatical change.
- ▶ Model-theoretic work on semantic change has often had a similar shape to generative work on grammar change.
- ▶ This will lead on to questions about distinctive elements of semantic change *vis-à-vis* other types of grammatical change.

Reanalysis

- ▶ The central idea in theories of grammatical change: diachronic relations between grammars are indirect, and mediated by linguistic behaviour.
- ▶ The linguistic behaviour offers only a partial, opaque, and unreliable cue as to the nature of the grammar.
- ▶ The grammar is different in kind from the behaviour.
 - ▶ This is the real basis of the argument from the poverty of the stimulus: no amount of behaviour gets you a grammar.
- ▶ The child observes the behaviour and makes guesses about the grammar. The result is a new grammar.
- ▶ Applicable to phonology (Andersen 1973), syntax (Lightfoot 1979), semantics (Traugott & Dasher 2002, Eckardt 2006), and arguably to any other domain which implicates induction of rule systems from examples of the outputs of those rules.
- ▶ Recently rediscovered under the name of the *Iterated Learning Model* (Hurford 1989, Kirby 2001, Kirby & Hurford 2002, Griffiths & Kalish 2007, Kirby et al. 2008, Smith 2009, Reali & Griffiths 2009). Same thing.

Reanalysis and the ILM

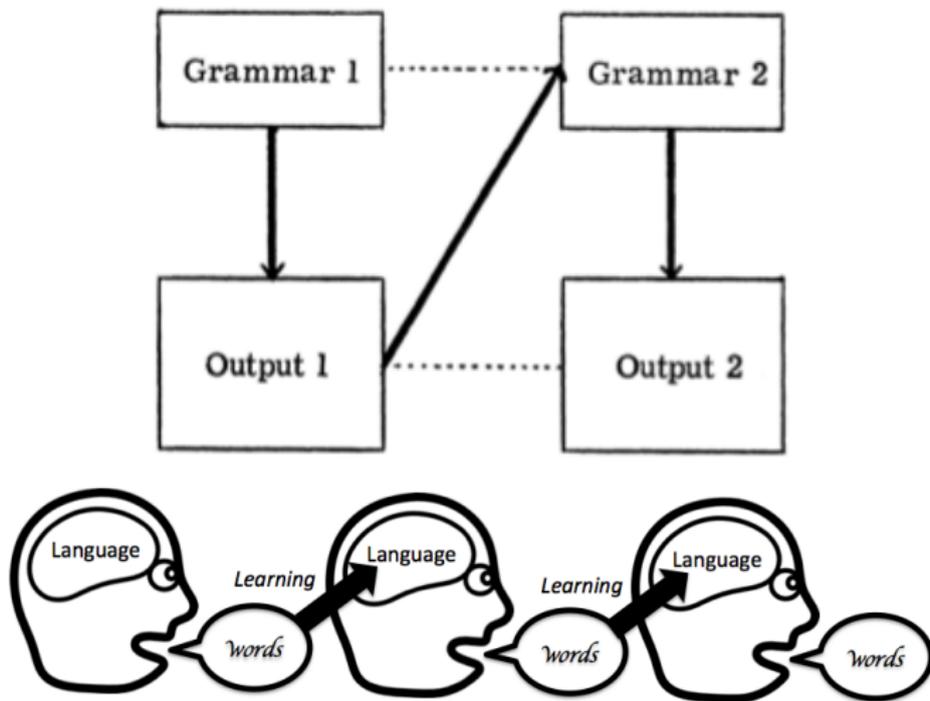


Figure 1 : Grammatical change mediated by linguistic output (Andersen 1973: 767 and *Replicated Typo* blog)

Reanalysis is everywhere, and hidden

- ▶ Reanalysis is just analysis. The grammar does not inhere in the data.
- ▶ There is a many–one relationship between grammars and finite outputs, and all we will ever experience (as analysts or learners) is a finite output.
- ▶ Extensionally, the following two scenarios are the same.
 1. A model uses grammar G_i to generate output $\{S_1, \dots, S_n\}$. A learner induces G_i on the basis of that output.
 2. A model uses grammar G_i to generate output $\{S_1, \dots, S_n\}$. A learner induces $G_{j \neq i}$ on the basis of that output, but (as it happens) only produces G_i -compatible outputs using G_j .(Intensionally, of course, they are quite different, but most diachronic work is doomed to be extensional.)
- ▶ Conclusion: reanalysis may occur without any detectable consequences.

There is no single target grammar

- ▶ Further conclusion: There is no single target grammar that a learner is meant to induce.
- ▶ Language identification in the limit (Gold 1967) is largely about target grammar identification, but it is not at all clear that target grammar identification is a naturalistic way to think about the challenge of language acquisition.
- ▶ It is not particularly interesting whether a learner acquires the putative target grammar (no gold star for acquiring G_i). More interesting are the range of types of mislearning, and implications for variation within a population.
- ▶ Some (“vacuous”) mislearning may have no empirical consequences. You were meant to learn one grammar; you learnt a different grammar which you use to do the same thing.
- ▶ Other types are diachronically more interesting.

Subset/superset mislearning

- ▶ This is also ubiquitous.
 - ▶ Imagine that my grammar of English contains the adjective *dreich* with a given meaning.
 - ▶ You learn from me but never learn *dreich*.
 - ▶ Your grammar generates a proper subset of the strings of mine.
- ▶ Same happens if you learn a word that I don't have. Again, ubiquitous.
- ▶ Same thing happens less trivially with little bits of structure (e.g. *notwithstanding* as a postposition).
- ▶ Less trivially still, your grammar shouldn't be massively more expressive than mine (the role of the Subset Principle).
- ▶ But as a matter of fact, slightly more expressive grammars (compared to the models) are learned all the time. New uses for *wh*-pronouns considered in Lecture 3 are one example.

“Probably Approximately Correct” mislearning

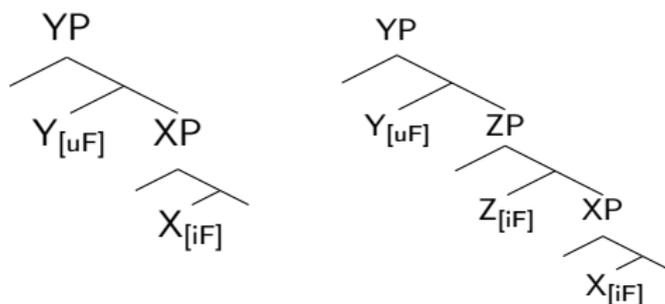
- ▶ The idea that identification of a target grammar is not necessarily the ideal model of grammar induction goes back at least to the “probably approximately correct” paradigm (Valiant 1984) — see Niyogi (2006) for a discussion of major results and their relevance to grammar learning.
- ▶ PAC grammar induction is successful if a learner selects a grammar that, with high probability, has low error in generalizing beyond the input given. Sounds intuitively like what a language learner should be trying to do.
- ▶ Unfortunately, PAC learning cannot even learn an arbitrary finite language (Niyogi 2006).
- ▶ So in terms of learnability in the “language identification in the limit” sense, PAC learning doesn’t get us any closer to an account of natural language learning.

Non-random near-misses

- ▶ Nevertheless, the idea that learners learn something in the right ballpark is a conceptually attractive one.
- ▶ Part of the problem might be in the specification of the question
- ▶ A learner's language must be close enough to his models to allow reasonably successful communication, but needn't be identical.
- ▶ Learners' near misses can be leveraged into a viable theory of change if they don't cancel each other out, because we've seen that change is likely to proceed along recurring pathways.
- ▶ Something must make certain near-misses more likely than others.

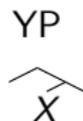
An example from syntax

- ▶ Roberts & Roussou (2003) on grammaticalization: a tendency for elements to be merged higher up the tree.
 - ▶ Stage 1: X is merged low but typically moved up the tree.



This gives a latent near-ambiguity (substantial overlap) between a grammar that moves X to Y and a grammar that merges X directly in Y .

- ▶ Stage 2: The item is merged higher.



Possible consequences: more “functional” semantics, knock-on effects on grammar of Y , etc.

An example from syntax

- ▶ Why would this latent near-ambiguity develop into an actual change?
- ▶ Roberts & Roussou's answer: the new grammar is more economical (e.g. preference for Merge over Move).
- ▶ Less controversial answer: something makes the new grammar more highly ranked according to some metric.
- ▶ Call that something *bias*. Bias could reflect several things:
 - ▶ Expectations during acquisition.
 - ▶ Production / usefulness / functional considerations.
 - ▶ Frequency.
 - ▶ Extent to which different grammars are supported by data.

Questions

- ▶ Near-miss learning raises a bunch of questions, some for a theory of grammatical change in general, some for a theory of semantic change in particular, all aimed at fleshing out the vague terms used above.
 - ▶ How does a latent (near-)ambiguity become an actual ambiguity?
 - ▶ How are ambiguities lost?
 - ▶ What counts as near?
 - ▶ What counts as a miss?
- ▶ The first two questions are quite general; answers to the last two questions will be specific to semantics.
- ▶ However, to my knowledge, there are currently no general answers to any of these questions.

From (re)analysis to change

- ▶ The external manifestations of a change are more visible than the change itself, but surely postdate the change (related to the actuation problem of Weinreich et al. 1968).
- ▶ Latent ambiguity is a prerequisite for reanalysis. It is pervasive (as guaranteed by a many-one mapping of grammars onto strings).
 - ▶ Grammar G_0 associates form F with structure S_0 .
 - ▶ Grammar $G_{\{0,1\}}$ associates F with structures S_0 and S_1 .
 - ▶ Grammar G_1 associates F with S_1 ($\neq S_0$) only.
- ▶ F is only ambiguous under $G_{\{0,1\}}$, but a learner exposed to F may induce any of these grammars: F underdetermines choice of G .
- ▶ Ambiguity arises (at a population level) when a first speaker induces $G_{\{0,1\}}$ or G_1 instead of G_0 . The ambiguity disappears when everyone has G_1 .

From (re)analysis to change

- ▶ Quite likely, most “changes” involve isolated speakers inducing $G_{\{0,1\}}$ or G_1 and have no population-level effect at all. This is similar to Weinreich et al.’s (1968) “base level of fluctuation” in a speech community.
- ▶ A tension is emerging:
 - ▶ Change at the population level is an aggregate of individual (re)analyses, which are not individually constitutive of change and may not even be individually observable.
 - ▶ But strictly speaking, those individual (re)analyses are all there is.
- ▶ Missing pieces:
 - ▶ How can individual analyses meaningfully differ without intolerable effects on intelligibility (today)?
 - ▶ How do the individual grammars relate to population-level “language” (tomorrow)?

Semantic near-misses

- ▶ A few claims, none of them very controversial:
 1. Utterance meaning is not equivalent to truth conditions, but closer to something like communicative intention.
 2. Compositionally derived sentence meaning is one clue to communicative intention, but doesn't exhaust the evidence marshalled in determining communicative intention.
 3. Compositionally derived sentence meaning is itself a multifaceted object (entailments, implicatures, presuppositions, etc.).
- ▶ A near miss could involve:
 - ▶ drawing the line between compositional and “ampliative” interpretation in a nonstandard place;
 - ▶ drawing the line between e.g. presupposition and assertion in a nonstandard place;
 - ▶ postulating a denotation that matches the “standard” denotation in most but not all cases;
 - ▶ Etc.

Potential theoretical implications

- ▶ The richer the array of semantic objects we work with, the more scope there is for theories of semantic reanalysis.
- ▶ Reanalysis depends on hidden structure to account for the many–one relation between grammars and stringsets.
- ▶ In principle, ability to account for diachronic phenomena could be a criterion in evaluating synchronic semantic theories. We're not there yet, though.

Eckardt on Jespersen's Cycle

- ▶ Eckardt (2006) identified a semantic counterpart of Jespersen's Cycle (normally conceived of as primarily syntactic) in French.
 1. Literal meaning of *pas/mie*, etc.
 2. NPI
 3. Negator
- ▶ Several variations on a theme.
 1. I didn't eat *mie* ('crumb')
 2. "I didn't eat even a small amount of food"
 3. "I didn't participate in even a minimal food-eating event"
 4. → *mie* as marker of minimal events.

The NPI stage

- (1) Por vos sui en prison mise . . . , Mais . . . longement n'i
for you I-am in prison put but for-long not-there
serai prise, se jel puis **mie**, . . .
I-will-be taken if I-it can 'mie'
"For you I am put in prison, but not for long will I be kept there, if I
can do anything about it"
- (2) le plus vaillant, le plus entier c'on trovast **mie** en tout le
the most brave the most whole that-one find 'mie' in all the
monde
world
"the most brave, the most accomplished that one finds ever in the
whole world"
- (3) Öistes vos s'il vendra **mie**?
Heard you if-he will-come 'mie'
"Did you hear if he will ever come?"
- ▶ This is semanticization of an invited implicature: something which could be inferred comes to be part of the denotation of *mie*.

Condoravdi & Deo on the perfect and perfective

- ▶ Indo-Aryan *-ta* generalized its meaning in three stages:
 1. Resultative
 2. Perfect
 3. Perfective

There's also a stable stative *-ta* that we won't discuss.

Diagnostics of progression beyond resultatives

- ▶ Resultative meanings require change-of-state predicates (distinguished from stative predicates by usual criteria like agents, instruments).

(4) yád áyā-taṃ dívodās-āya vartí-ḥ. . .
when come-impf.2.du D-dat.sg abode-acc.sg
revád uvāh-a sacan-ó rátho
riches.acc.sg carry-pfct.3.sc good-m.pl chariot.nom.sg
vāṃ vṛṣabhá-ś ca śiṃsumāra-ś ca
you.gen.du bull-nom.sg and dolphin-nom.sg and
yuk-tā

yoke-perf.m.pl

“When you (Aśvins) *came* to Divodāsa, (to his) abode, your chariot *carried* rich goods. A bull and a river dolphin *were yoked* to it.”

Diagnostics of progression beyond resultatives

- ▶ This means that *-ta* forms with other predicates are diagnostic of generalization beyond resultative meanings.

(5) a. **Existential perfect**

mantrā nānāprakār-āḥ sy-ur
formula.nom.m.pl various.sort-nom.m.pl be-opt.3.pl

dṛṣ-ṭā ye mantridarśi-bhiḥ
see-perf.m.pl which.nom.pl seer-ins.pl

“The formulas, which have been seen by the sages (or seers), may be of various sorts”

b. **Universal perfect**

ta-smin sarva-ṃ **pratiṣṭh-itaṃ** yat ca
it-loc all-nom.sg rest-perf.n.sg which and

prāṇi yat ca na
live-pers.3.sg which and neg

“On it (milk) everything has rested; that which lives and that which does not.”

Evidence for emergence of the perfective

- ▶ Evidence for progression from perfect to perfective meaning includes use to advance reference time of narrative, and use with past time adverbials.

(6) **tato** kaiva-esu divas-esu aikkan-t-esu...
then many-loc.pl day-loc.pl pass-perf-loc.pl
diṭ-ṭhā me taruṇajuvati
see-perf.f.sg I-ins young.woman.nom.f.sg
“Then, upon the passing of many days, I saw the young woman.”

Generalization in model-theoretic terms

- ▶ Core of Condoravdi & Deo's analysis: the denotation of *-ta* at stage i entails the denotation at stage $i + k$. This is what generalization looks like model-theoretically.

(7) a. **Resultative**

$$\lambda\langle P, Q \rangle \lambda i \exists e, s [P(e) \wedge Q(s) \wedge \mathbf{result}(e, s) \wedge i = \tau(s)]$$

b. **Perfect**

$$\lambda P \lambda i \exists e, j, k [P(e) \wedge \tau(e) \sqsubseteq k \wedge i, k \sqsubseteq j \wedge k \prec i]$$

c. **Perfective**

$$\lambda P \lambda i \exists e, j [P(e) \wedge \tau(e) \sqsubseteq i \wedge i \sqsubseteq_{\text{final}} j]$$

- ▶ This diachronic progression is quite standard (also — partially at least — in English, French, etc.).
- ▶ It illustrates meaning change in functional elements rather than recruitment of lexical elements to functional uses.
- ▶ Condoravdi & Deo suggest that this may be why their analysis relies on mislearning rather than semanticization of invited inferences.

Discussion

- ▶ Even Eckardt's NPI story requires mislearning, though.
 - ▶ The availability of things like *mie* as NPIs is mundane and probably universal.
 - ▶ What's interesting is a putative diachronic shift from *mie* as predicate of individuals ($\lambda x.\text{crumb}(x)$) to *mie* as predicate of events ($\lambda P_{\langle Ev, t \rangle} \lambda e.P(e) \wedge \neg \exists e' < e.P(e')$).
 - ▶ Most crumb-manipulating events are minimal events, so there's a near-miss entailment relation between the two *crumbs*.
- ▶ The diachrony of the perfect shows a steady accumulation of denotations: resultative and stative readings are available throughout, so new perfect and perfective readings do not eliminate the older readings.
- ▶ The model-theoretic formulation helps explain the implicational pattern (you wouldn't expect resultative + perfective without existential/universal perfect on this analysis), and the co-occurrence of existential and universal perfect readings.
 - ▶ The above denotation gives you both together.
 - ▶ To exclude one you'd need a more complex denotation which probably wouldn't be entailed by the resultative denotation.

Cournane on parallel pathways in acquisition and change

- ▶ Cournane (2015): acquisition pathways can align with diachronic pathways.
 - ▶ Root modal meanings acquired before epistemic.
 - ▶ Root modal lexemes develop epistemic uses diachronically.

- (8)
 - a. he mus(t) talk [Deontic; Sarah 2;9, referring to her broken talking-doll]
 - b. must be gone [Epistemic; Sarah 3;0, referring to missing toy plates]

- (9)
 - a. Tie minnera habeton die **muosan** gan
Those no.money had they had.to walk
“those who had no money were obliged to walk”
MHG, c.1000
 - b. Du **must** wohl müde sein
you must mod.part tired be
“You must be tired”
Modern German

- ▶ Cournane relates parallel pathways to overextension. Children need to generalize beyond the input given, and generalizing too much in a given direction gives a new meaning

Cournane on opposing pathways

- ▶ Syntactically, acquisition of modals proceeds in the opposite direction to diachrony
 - ▶ Monoclausal structures acquired before biclausal.
 - ▶ Biclausal premodal structures develop into monoclausal modal structures.

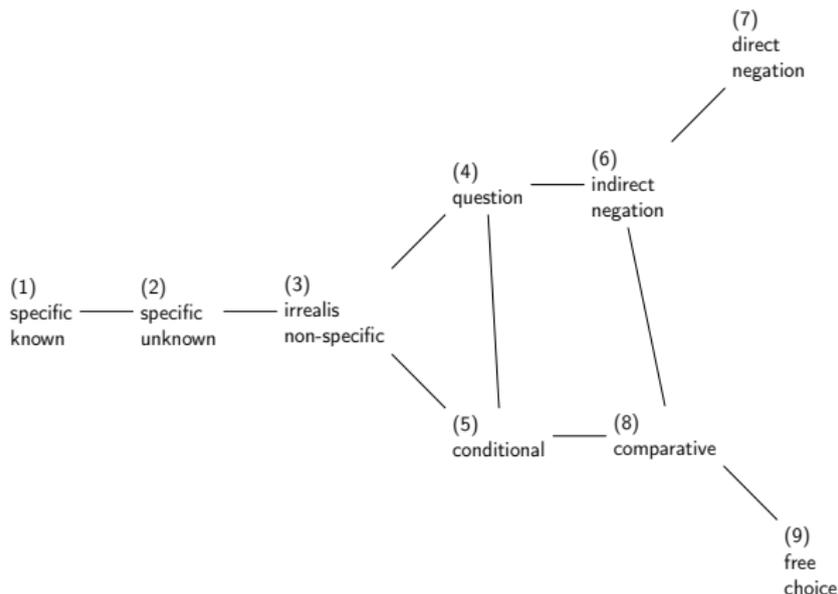
- (10) a. want Bobo [2;3,07]
b. wan go read it [2;7,28]
c. I want Daddy to help me [3;3,13]

- (11) a. sone hit **mæi** illimp-en
soon it may.3sg happen.inf
“It may happen soon” ME, c.1200
b. It may happen soon.

- ▶ Cournane's explanation: modal syntax arises when children fail to go beyond a preliminary analysis for a verb like *magan*, so the diachronically later syntax matches an acquisitionally earlier analysis.

Semantic maps

- ▶ Many of the above ideas can be related to the notion of semantic map. E.g. Haspelmath (1997) on indefinite pronouns.



- ▶ Lexical items can be used for a continuous subset of this map (e.g. English *any* can be used for 4–9, Dutch *enig* for 4–6, 8).

Semantic maps and semantic change

- ▶ Semantic change can be understood as annexing or relinquishing space on the map; the map itself represents constraints on possible changes, as well as constraints on possible languages: a panchronic typology.
- ▶ Those constraints (like the maps themselves) can be understood in two ways: either the analyst starts with a dataset and induces the map based on usage of individual lexical items (the *matrix-driven* approach of Zwarts 2010) or starts with a theory of some semantic space and maps lexical items on to it (the *space-driven* approach).
- ▶ The above three examples all imply small parts of semantic maps.

- (12)
- a. Small entity — Minimal entity/event (NPI) — Negation
 - b. Stative — Resultative — Perfect — Perfective
 - c. Deontic — Epistemic

Semantic maps and semantic change

- ▶ Ideally, we will ground many aspects of semantic maps in our model theory.
 - ▶ This is what Eckardt, or Condoravdi & Deo, are attempting.
 - ▶ If denotation D_1 entails D_2 , and D_2 is the sort of thing that a functional item could denote, then that grounds a connection between D_1 and D_2 in a semantic map.
 - ▶ Possible word meanings are then D_1 , $D_2 (= D_1 + D_2)$, possibly $D_2 \setminus D_1$?
- ▶ Such reasoning may not always work, though.
 - ▶ NP-intensifier *self* can develop into a clausal focus particle (German *selbst*) and a reflexive marker (PDE *-self*).
 - ▶ You might be able to tell some story about entailment for the German case (though the fine detail of Eckardt's account suggests not).
 - ▶ But I struggle to see how there can be an entailment relation between NP-intensifier *self* and reflexive *-self*.

Semantic maps and semantic change

- ▶ In such cases, “semantic maps” possibly don’t represent neat facts about possible or probable denotations (see also Zwarts 2010 on discontinuous denotations in semantic maps), and the English *self* case may be actual lexical ambiguity.
- ▶ Corollary: The space-based approach to semantic maps won’t generate a complete typology.
- ▶ So both methodologies have their merits: the space-based approach is more explanatory; the matrix-based approach is more complete.
- ▶ *cf.* Gärdenfors 2014: “The different semantic maps should be connected and unified so that we generate a picture of the entire semantic universe—the geography of the human mind” (!).

Semantic maps and semantic change

- ▶ The maps given above are not complete, or encapsulated (e.g. the NPI part of (12a) intersects with Haspelmath's 4–8). The diachrony of *wh*-forms sketched yesterday also interacts with Haspelmath's map.



- ▶ Also preliminary in many respects:
 - ▶ Potentially lots of detail in NPI-like “restricted indefinite” uses.
 - ▶ Not clear that the free → appositive → restrictive pathway can all be reduced to the semantics of the *wh*-form.
 - ▶ The original interrogative / indefinite / correlative triangle could use clarification.
- ▶ Each individual *wh*-form is following its own path through the map (the lexical differences mentioned yesterday).

Comments

- ▶ The diachrony of *wh*-forms is broader than grammaticalization — in particular, no necessary link to bleaching or generalization (or correlates in syntax and phonology).
- ▶ In fact, not clear that expressivity, entailment, etc. will have any explanatory value here.
- ▶ Perhaps the development of restrictive relatives reflects change elsewhere in English. Free relatives are definite descriptions. These can be derived compositionally in at least two ways:
 1. $[\text{Wh NP}] = \lambda s. \iota x. [\text{NP}](x)(s)$ (e.g. Jacobson 1995)
 2. $[\text{Wh NP}] = \lambda x. [\text{NP}](x)$
 $[\delta \text{ wh NP}] = \lambda s. \iota x. [\text{NP}](x)(s)$ (e.g. Caponigro 2003)
- ▶ 1. is simpler. 2. is more flexible (allows indefinite FRs and restrictive RCs).
- ▶ OE didn't have a definite article. *The* emerged in ME.
- ▶ Perhaps reanalysis of type-1 *wh* to type-2 *wh* took place, related to the broader reorganization of definiteness marking.

What drives change relative to semantic maps?

- ▶ We're accumulating quite a list of motivations for semantic change:
 - ▶ Expressivity
 - ▶ Entailment/generalization/bleaching
 - ▶ Mislearning
 - ▶ ...
 - ▶ In itself, that's no bad thing, especially as the different factors potentially have different properties.
 - ▶ Mislearning lends itself most readily to child-driven theories of change;
 - ▶ Expressivity is relevant throughout one's lifetime (Eckardt says particularly in adulthood).
 - ▶ Still, there are redundancies here. Take generalization.
- Case 1:** The target grammar for expression E has denotation D_1 , the learner has D_2 , where $D_1 \subset D_2$. This is mislearning.
- Case 2:** The learner also has D_1 but nevertheless uses E for $D_2 \setminus D_1$. This is expressivity.
- ▶ So generalization plausibly reduces to mislearning and/or expressivity. But each of those has its own problems.

What drives change relative to semantic maps?

- ▶ Mislarning: We learn some pretty detailed lexical entries. That might suggest that mislarning is isolated and so unlikely to lead to change (Weinreich et al.'s transmission problem).
- ▶ Expressivity: Only plausible if:
 1. there's no better way to express $D_2 \setminus D_1$ (otherwise pragmatics should make the listener wonder why the speaker's being weird);
 2. the desire to express $D_2 \setminus D_1$ arises so rarely that use of E to express this meaning hasn't already been conventionalized.

E.g. in Eckardt's analysis of Jespersen's Cycle:

- ▶ Minimal elements get used as NPIs all the time. That's unlikely to be an expressivity-based change.
- ▶ Hard to know what conventionalization would look like as a change (probably not a semantic change).
- ▶ There were a range of semantically similar (though admittedly nonidentical) minimizers. They can't *all* have emerged (and been maintained) for reasons of expressivity.

Summary

- ▶ Semantic change is grammar change.
- ▶ Grammar change is change in the properties of lexical items.
- ▶ Semantic maps can be a convenient proxy for “the properties of lexical items”, though we assume that such maps ultimately have no theoretical status.
- ▶ Much of grammar change comes down to types of mislearning, which drive interesting relationships between acquisition pathways and diachronic pathways.
- ▶ Other types of explanations have also been advanced (expressivity, generalization) — tomorrow will partly be about situating such explanations within a discussion of the population-level dynamics of change.

Bibliography I

- Andersen, H. (1973). Abductive and deductive change. *Language*, 49, 765–793.
- Caponigro, I. (2003). *Free not to Ask: On the Semantics of Free Relatives and Wh-words Cross-linguistically*. PhD thesis, University of California, Los Angeles, CA.
- Condoravdi, C. & Deo, A. (2015). Aspect shifts in Indo-Aryan and trajectories of semantic change. In C. Gianollo, A. Jäger, & D. Penka (Eds.), *Language Change at the Syntax–Semantics Interface* (pp. 261–292). Berlin: de Gruyter.
- Cournane, A. (2015). In defense of the child innovator. Ms., University of Toronto.
- Eckardt, R. (2006). *Meaning Change in Grammaticalization: An Enquiry into Semantic Reanalysis*. Oxford: Oxford University Press.
- Gärdenfors, P. (2014). *The Geometry of Meaning: Semantics based on Conceptual Spaces*. Cambridge, MA: MIT Press.
- Gold, E. M. (1967). Language identification in the limit. *Information and Control*, 10, 447–474.
- Griffiths, T. & Kalish, M. (2007). Language evolution by iterated learning with Bayesian agents. *Cognitive Science*, 31, 441–480.
- Haspelmath, M. (1997). *Indefinite Pronouns*. Oxford: Oxford University Press.
- Hurford, J. (1989). Biological evolution of the Saussurean sign as a component of the language acquisition device. *Lingua*, 77, 187–222.
- Jacobson, P. (1995). On the quantificational force of English free relatives. In E. Bach, E. Jelinek, A. Kratzer, & B. Partee (Eds.), *Quantification in Natural Languages* (pp. 451–486). Dordrecht: Kluwer.
- Kirby, S. (2001). Spontaneous evolution of linguistic structure—an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, 5, 102–110.
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, 105, 10681–10686.

Bibliography II

- Kirby, S. & Hurford, J. (2002). The emergence of linguistic structure: An overview of the iterated learning model. In A. Cangelosi & D. Parisi (Eds.), *Simulating the Evolution of Language* (pp. 121–148). Dordrecht: Springer.
- Lightfoot, D. (1979). *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
- Niyogi, P. (2006). *The Computational Nature of Language Learning and Evolution*. Cambridge, MA: MIT Press.
- Reali, F. & Griffiths, T. (2009). The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition*, 111, 317–328.
- Roberts, I. & Roussou, A. (2003). *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge University Press.
- Smith, K. (2009). Iterated learning in populations of Bayesian agents. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, (pp. 697–702).
- Traugott, E. & Dasher, R. (2002). *Regularity in Semantic Change*. Cambridge: Cambridge University Press.
- Valiant, L. (1984). A theory of the learnable. *Communications of the ACM*, 27, 1134–1142.
- Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. Lehmann & Y. Malkiel (Eds.), *Directions for Historical Linguistics* (pp. 95–188). Austin, TX: University of Texas Press.
- Zwarts, J. (2010). Semantic map geometry: Two approaches. *Linguistic Discovery*, 8, 377–395.