

Modeling Contributions of Usage versus Acquisition to Language Change

Andy Wedel
University of Arizona

Clay Beckner
University of New Mexico

Two mechanisms of change

- Acquisitional change: Error/innovation occurs when L1 learners deduce the wrong grammar from adult utterances.
- Usage-based change: Innovation occurs as a result of the act of communication. e.g., spread or loss of variants due to frequency effects, analogy, social prestige.

Two mechanisms of change

There is a long history of attributing change to either acquisition, or usage...

(See Jespersen 1922)

Language change originates in imperfect learning:

‘The chief cause of sound change lies in the transmission of sounds to new individuals’ – Paul (1880)

‘All the major changes in pronunciation that we have been able to investigate originate in child speech’ – Passy (1891)

Language change originates in imperfect learning:

‘If languages were learnt perfectly by the children of each generation, then languages would not change.’ –Sweet (1899)

‘A basic cause of change is the way children acquire the language. . . The child’s grammar is never exactly like that of the adult community’ – Fromkin & Rodman (1993).

See also Halle (1962), Andersen (1978), Lightfoot (1991, 2006), Niyogi & Berwick (1995), Newmeyer (1998).

Language change originates in usage, not transmission:

Children's acquisition is 'in most cases practically perfect...the main cause of sound-change must therefore be sought elsewhere...[in] organic shifting, failure to hit the mark, the result either of carelessness or sloth' –Sweet (1900)

'In historical change of existing languages, it appears that lasting innovations do not come from preschoolers but from older speakers. That is, language changes more *in use* than it does in the process of being learned.' –Slobin (2004)

'To summarize, babies do not initiate changes. Groups of interacting speakers do, particularly adolescents.' –Aitchison (2001: 739)

See also Drachman (1978), Barlow & Kemmer (2000), Croft (2000)

Why believe acquisition drives change?

- A principal motivation is the belief that L1 learners are the only possible source of change, due to critical period effects.
- Strong claim: Grammar is unchangeable past a critical period, and adults can only adjust their utterances. (Halle 1962, Newmeyer 1998, Senghas et al. 2004, Lightfoot 2006)

Why believe acquisition drives change?

- Weaker claim: Young speakers learn/change more easily than older speakers. (Newport 1990)
- Past sensitive period, adults are less able to learn new languages (or innovations) fluently.

Language learning declines with age of exposure

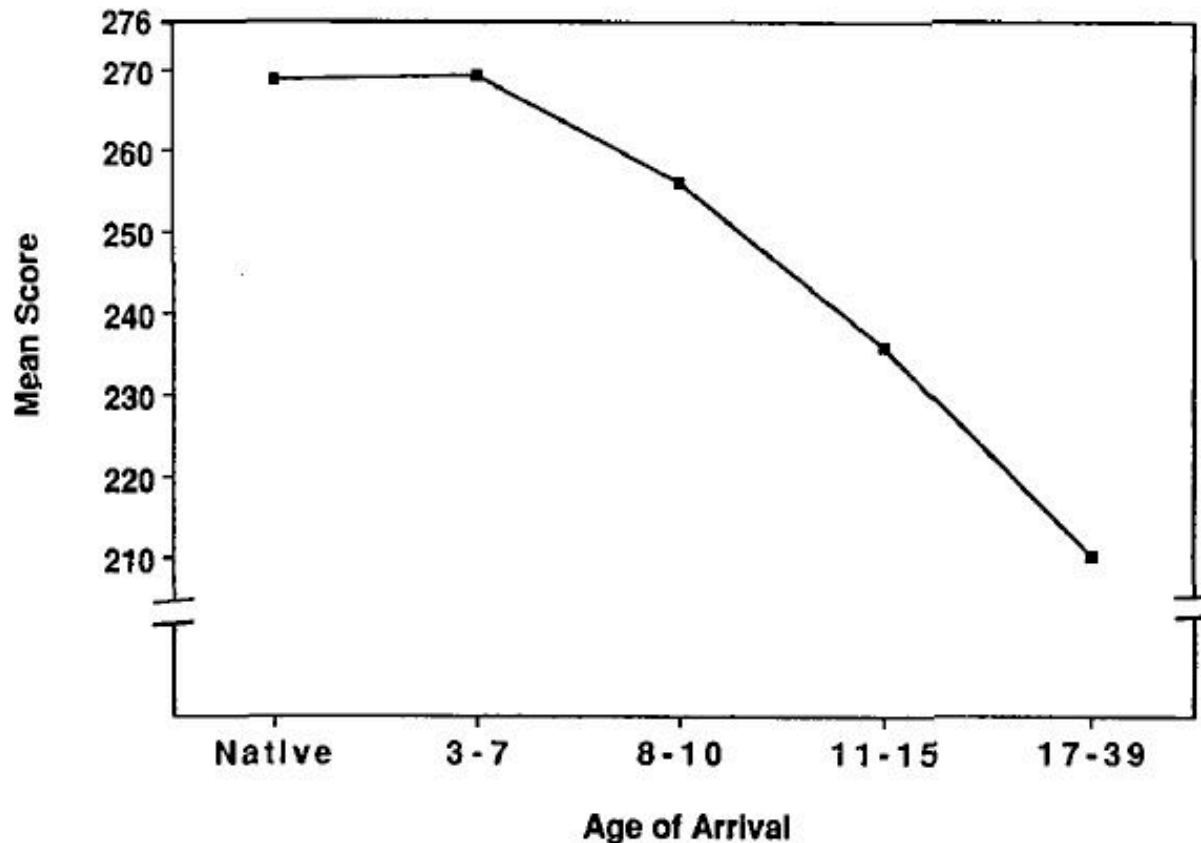


Figure 2. Total score on a test of English grammar in relation to age of arrival in the United States (redrawn from Johnson & Newport, 1989).

Why believe usage drives change?

1. Change can occur only when innovations persist and spread.
 - Successful change originates in socially influential groups, and L1 learners do not meet this criterion (Kerswill 1996, Aitchison 2001)
 - L1 learners copy adult speech (not vice-versa).
 - Thus innovations introduced by child L1 learners are usually transitory.
 - Many L1 innovations do not correspond to historical change (e.g., Drachman 1978, Vihman 1980, Croft 2000).

Why believe usage drives change?

2. Some changes are unlikely to be innovated via acquisition.

A. Reductive change.

- Frequent words and phrases undergo more phonetic reduction (Bybee 2007). These reduced forms may become conventionalized as part of the grammar (e.g., *going to* > *gonna*).
- The frequency correlation means that innovation increases with more L1 exposure. An imperfect learning model cannot explain change that increases on items with more training.

Why believe usage drives change?

2. Some changes are unlikely to be innovated via vertical transmission.

A. Reductive change (cont'd).

- Reduction does not make the language easier to acquire. e.g., *didn't* is harder for children to say, and it is more opaque than *did not* (Bybee to appear)
- Reductive innovations occur as a result of processing, and expert fluency arising from repetition. These are usage-based factors.

Why believe usage drives change?

2. Some changes are unlikely to be innovated via vertical transmission.

B. Maturation constraints

- Do adult speakers communicate using only the grammatical conventions created by L1 learners?
- Adults 'have more communicative needs than children to satisfy' and they face a broader set of domains in which they may need to innovate conventions (Mufwene 2001: 196).

Why believe usage drives change?

2. Some changes are unlikely to be innovated via vertical transmission.

B. Maturation constraints (cont'd)

- Some domains are cognitively inaccessible to children, and thus will not be innovated via imperfect learning (Slobin 1997, Bybee to appear).
- For instance, epistemicity is acquired late by children (Barbieri & Bascelli 2000), and it is thus unlikely that epistemic markers originate via L1 innovations.

Toward a synthesis

- First, we note that acquisition vs. usage is an artificial dichotomy.
- There is considerable evidence that adults learn continuously throughout the lifespan, in tandem with usage. e.g., sound change in adulthood (Sankoff 2004, Harrington 2006)
- Adult speakers can create/learn a (second) language through interaction, e.g., Tok Pisin (Sankoff & Laberge 1973, Romaine 1995)

Toward a synthesis

- Children can acquire a first language while the language is created via the acquirers' interactions. e.g., Nicaraguan Sign Language (Senghas 1995, Senghas et al. 2004)
- In sum, usage and acquisition are in many ways inseparable.

Toward a synthesis

- Many instances of change can reasonably be attributed to acquisition, usage, or both.
- For instance, both L1 learners and adults tend to regularize infrequent forms. (Bybee & Slobin 1982)
- Explanation: Infrequent irregulars aren't learned by children, and/or aren't accessible due to disuse.

Toward a synthesis

- Pinker (1999): '[This] is why irregular verbs tend to be high in frequency: the [lexicon] has been filtered repeatedly through the minds of children and adults, both of whom regularize uncommon irregular verbs.'

Toward a synthesis

- In a complex systems view, it is reasonable to expect that both usage and acquisition may contribute to change.
- It is also reasonable to expect that these mechanisms will interact with one another.

Toward a synthesis

- The present simulations are part of a larger project to investigate these interactions systematically.
- Simulations provide a methodology for investigating the roles of acquisition and usage, both separately and in interaction.

Change in a population requires turnover

- Can be thought of as information loss + information gain.
- Bottlenecks in biological evolution

Kirby: Language acquisition is an information bottleneck

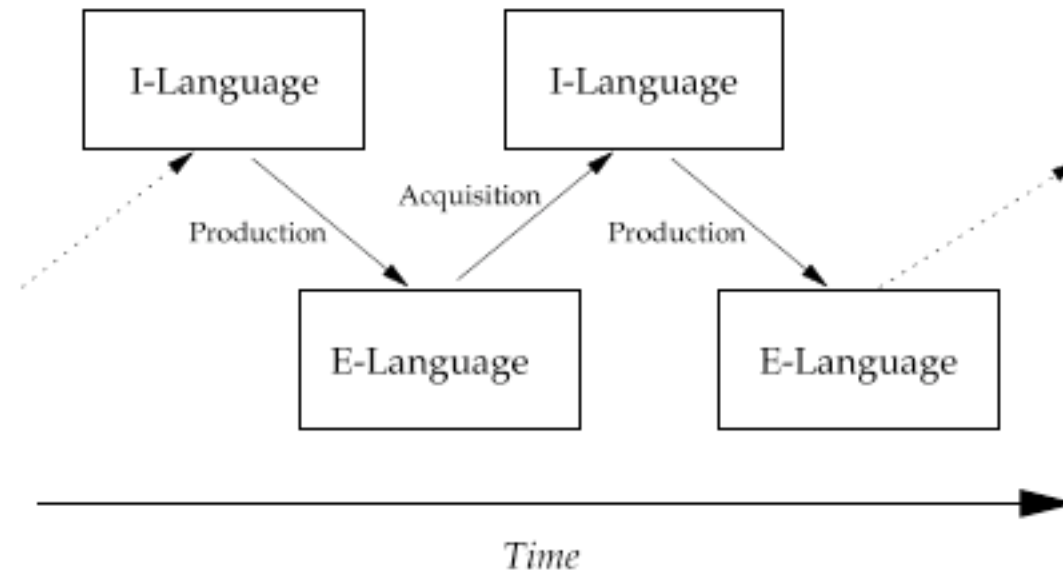
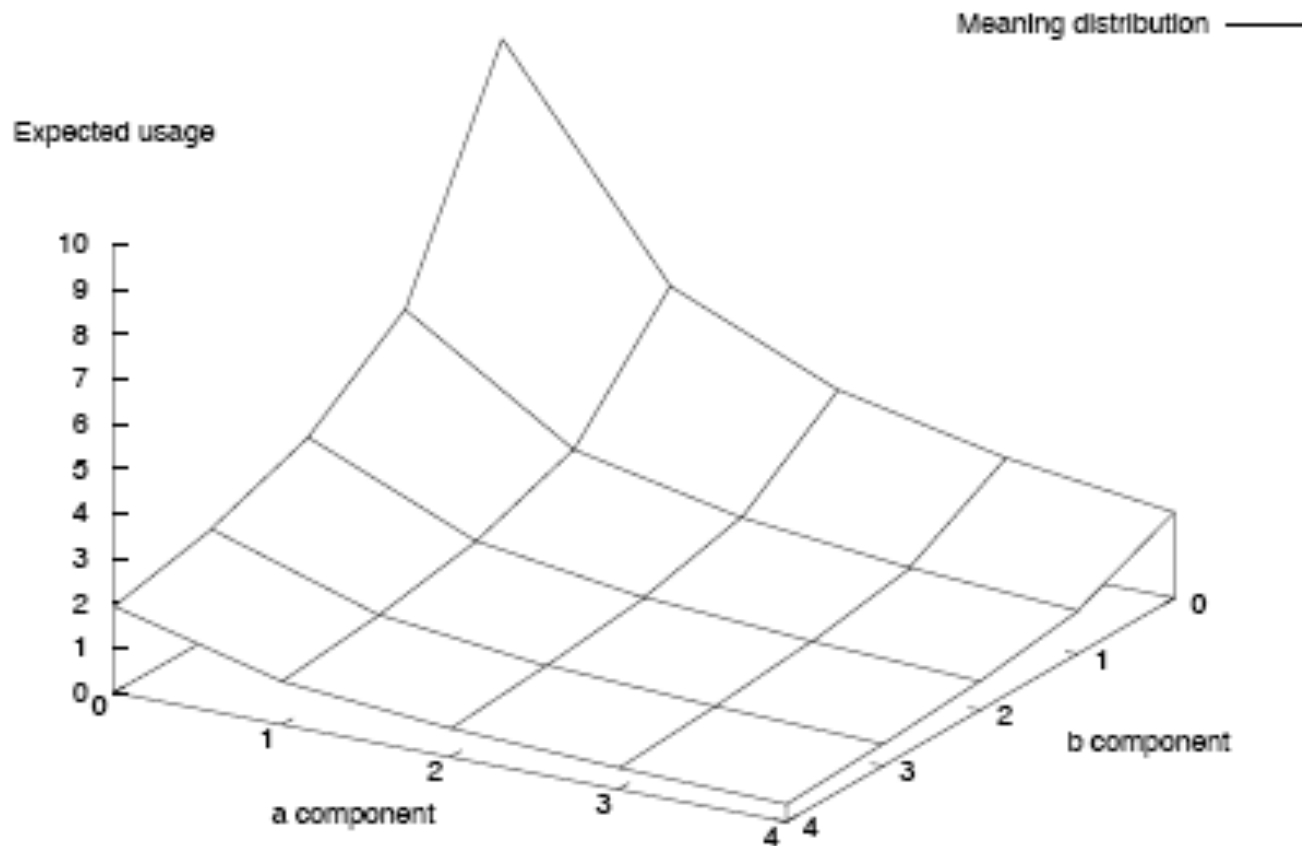


figure 1.1. The transmission of language over time.

Acquisition-based model of compositionality vs. idiosyncrasy

- 2 dimensional meaning space with 25 bimorphemic words
 - $a_{0-4} \times b_{0-4} = 25$ meanings
 - Each meaning is associated with a string of characters.
- Speaker utters 50 randomly chosen meanings to learner, who stores mappings between strings and meanings.
- Learner becomes new speaker, and utters another 50 random meanings to new learner.
 - If the meaning was encountered in the learning phase, the corresponding string is uttered.
 - If not, composition is attempted from other words sharing meaning components.
 - Small chance of deleting a character

Frequency differences of meaning expressions



Sample Results

	a_0	a_1	a_2	a_3	a_4
b_0	g	s	kf	jf	uhlf
b_1	y	jgi	ki	ji	uhli
b_2	yq	jgq	kq	jq	uhlq
b_3	ybq	jgbq	kbq	jbq	uhlbq
b_4	yuqeg	jguqeg	kuqeg	juqeg	uhluqeg

- Intended take home message: high frequency enables idiosyncrasy to survive the acquisition bottleneck.

Change through drift

- Information is lost and gained in any population with turnover, whether abrupt or gradual.
 - Drift in biological populations
 - Drift in populations within categories through usage
- Language change through information loss/gain could arise in principle in acquisition OR in usage.

Illustrate functional equivalence with Kirby-esque model

- Translate Kirby model into an exemplar format:
 - Words comprise 2 morphemes of 2 characters each.
 - {a1, a2, a3, a4} + {b1, b2, b3, b4}
 - Listener is privy to the speaker's meaning, and stores all percepts as exemplars of that meaning category
 - Generalizations about string-morpheme mappings are passively stored within the total exemplar set, rather than as explicit rules.

Illustrate functional equivalence with Kirby-esque model

- Translate Kirby model into an exemplar format:
 - Provide two pathways for production of a meaning:
 - Holistic: An exemplar is chosen at random from the set for the current meaning and produced.
 - Compositional: The most frequent character string types for each morpheme in the current meaning are concatenated and produced.
 - Bias choice of pathway by frequency
 - Highly frequent forms rarely use the compositional pathway.
 - Very infrequent forms use the compositional pathway relatively more.

Bottleneck versus Drift Models

- Bottleneck:
 - Two agents, conceived as teacher and learner.
 - Teacher produces 200 outputs for a naive learner, who stores each as an exemplar in the corresponding meaning category.
 - Learner becomes teacher to new naive learner.
 - Teacher's choice of holistic versus compositional routes for each production influenced by the number of times the teacher heard each form as a learner.

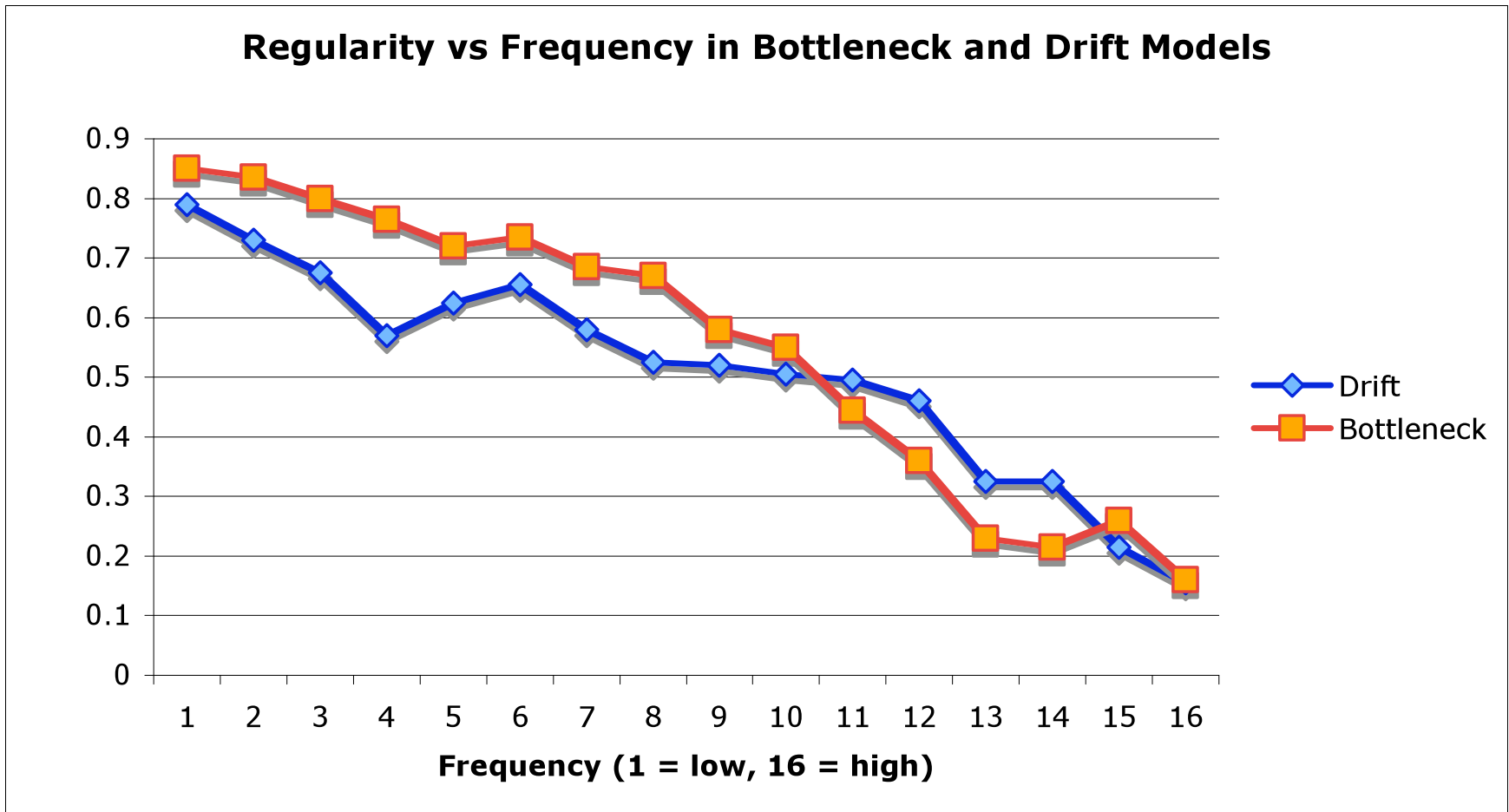
Bottleneck versus Drift Models

- Drift:
 - Two agents conceived as continuously conversing to each other.
 - Speaker produces an output which is stored as a new exemplar by the listener, displacing an old exemplar.
 - Speaker's choice of holistic versus compositional routes for each production influenced by the frequency with which that meaning has been previously heard.

Results

- Run 200 independently seeded simulation runs in both drift and bottleneck mode.
 - Each run involves 40,000 productions
- Calculate a regularity score for the most common exemplar in each meaning category
 - 1 if morphemes match the highest type frequency in the lexicon.
 - 0 otherwise.
- Average scores for each meaning over 200 runs.

Results



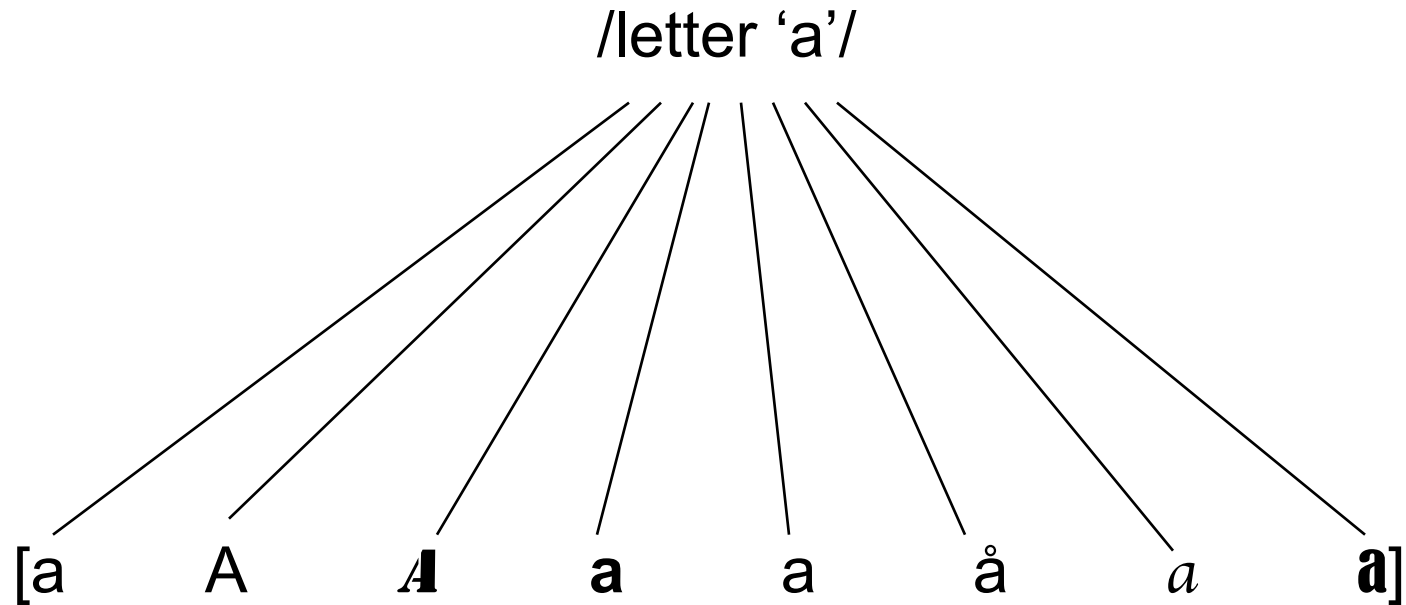
Interim conclusion

- Change through information turnover in ‘acquisition’ versus ‘usage’ is not qualitatively different at this level of abstraction.
- What *is* qualitatively different between acquisition and usage?
 - Rate of category label creation/loss

Category creation versus category usage

- Category creation involves abstraction of category labels from distributions in input data.
 - Language acquisition
- Usage involves mapping data to and from existing category labels.
 - Language use

Categories contain variants



Two qualitatively different kinds of change given this conception:
Change in category labels, or change in category contents.

Influence of existing categories on present behavior

- Wide array of evidence that perception and production are influenced by preexisting category structure.
 - Perception:
 - Categorical perception
 - Perceptual magnet effect (Kuhl 1991)
 - Production:
 - Motor patterns (e.g., Zanone & Kelso 1997)
 - Language change patterns (cf. 'Choice', Blevins 2004)

Category creation versus category usage

- Acquisition and usage contribute qualitatively distinct pathways of language change.
 - Acquisition
 - Changes in category system
 - Usage
 - Changes in category contents
- Each constrains the other.

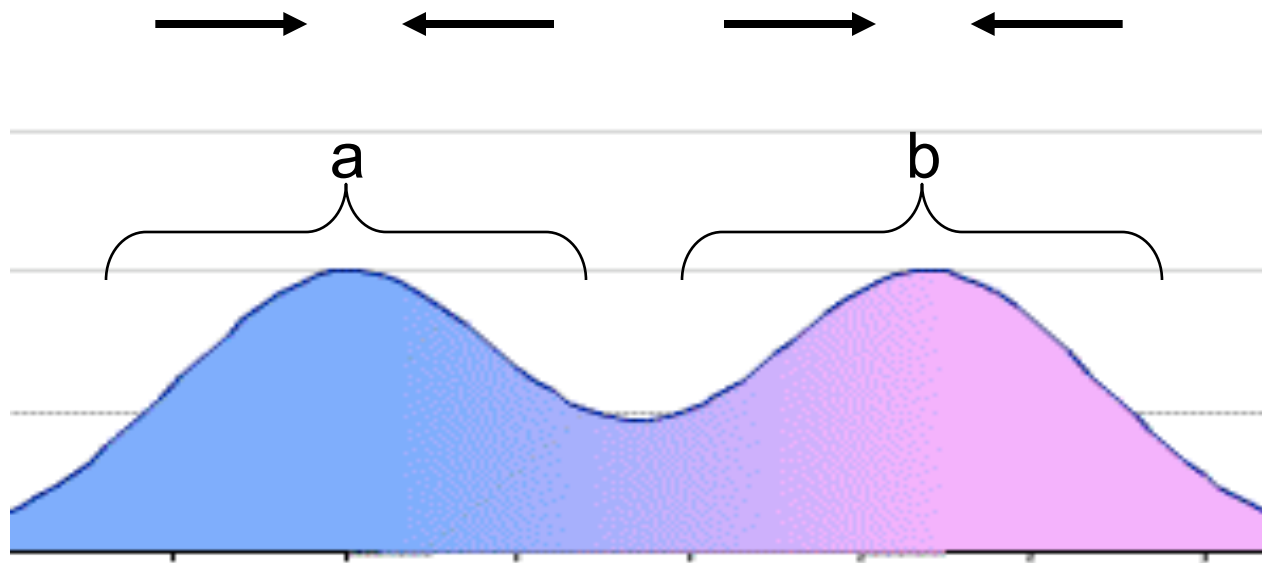
Illustrate with contrast-maintenance simulation

- Horizontal, usage phase
 - Agents produce forms with reference to existing sound categories
 - Sound outputs are an average, with sampling biased toward within-category exemplars.
- Vertical, acquisition phase
 - Agents abstract new sound categories on the basis of the distribution of perceived sounds.

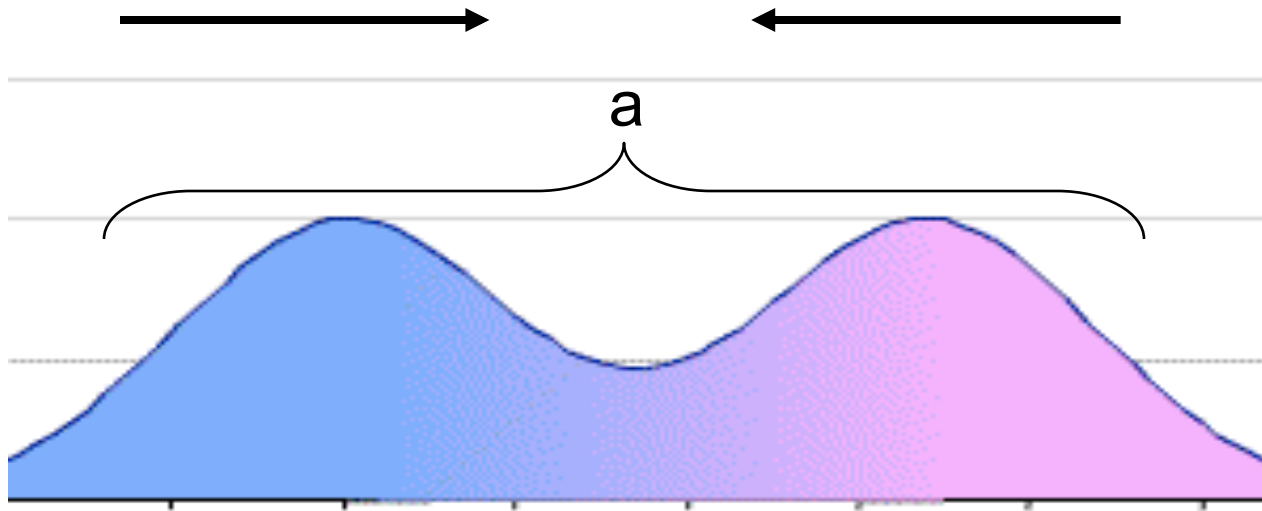
Expected properties of this architecture

- Change in category contents in the horizontal phase influenced by the existing sound category labels
- Change in the set of sound category labels influenced by changes in category contents.

Horizontal phase: two sound categories



Horizontal phase: one sound category



Vertical phase

- Depending on the breadth of a bimodal distribution from the preceding horizontal phase, it may be acquired as one or two categories in the vertical phase.
 - with distinct repercussions on the trajectory of change in the *following* horizontal phase.

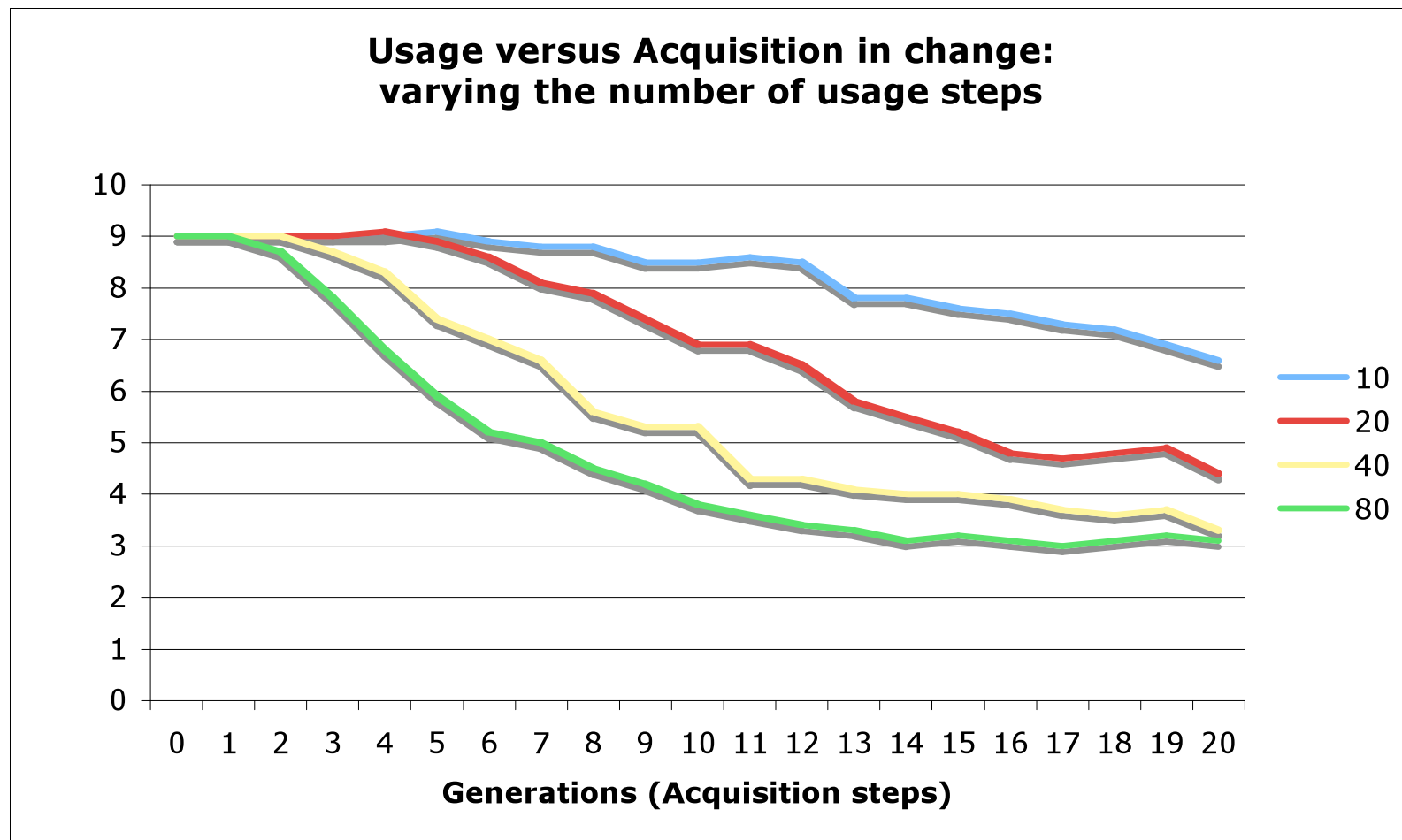
The interdependency of category label abstraction and category use

- Two experiments:
 1. Keep the number of total vertical acquisition steps constant, but vary the number of horizontal usage steps between.
 2. Keep the number of total horizontal usage steps constant, but vary the number of vertical acquisition steps.

Experiment structure

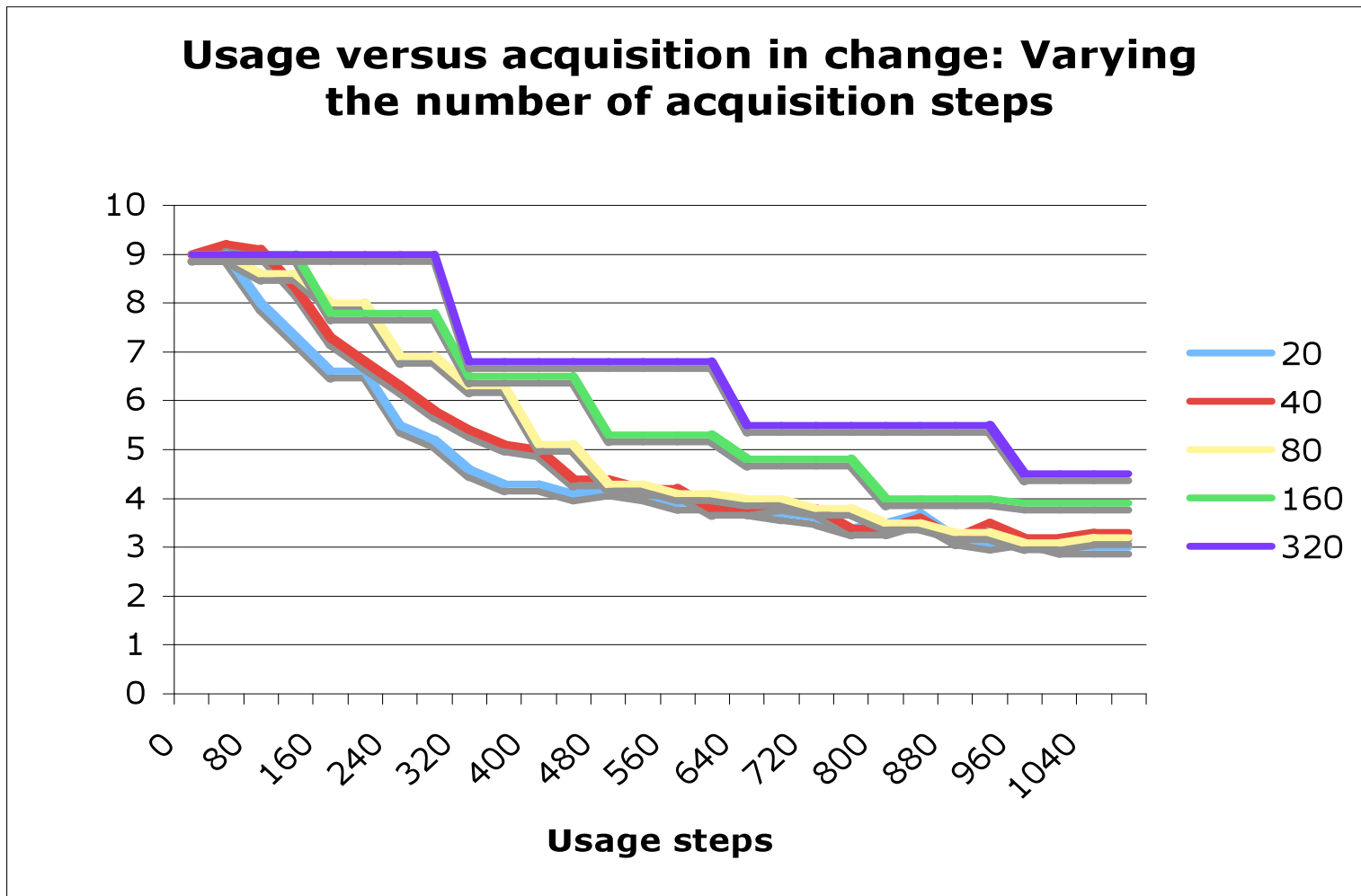
- Six word categories
- Seed with a sound distribution that will be initially acquired into nine categories.
 - With six word categories each comprising two sounds, three sound categories is the most stable state in this simulation architecture.
- Look at rate of relaxation from nine towards three sound categories.

1: Acquisition steps constant, usage steps vary



Each point is the average of 10 independent runs

2: Acquisition steps vary, usage steps constant



Each point is the average of 10 independent runs

Conclusions

- *Usage versus acquisition* may not be a particularly useful apposition in understanding language change given that they do not refer to well-distinguished mechanisms.
 - Acquisition involves usage
 - At least some kinds of language acquisition continue throughout life

Conclusions

- A more useful distinction may lie between change in *category labels* versus change in *category mappings*
 - Roughly correspond to frequent types of changes that occur in ‘acquisition’ versus ‘usage’.

Conclusions

- We illustrated interaction of both types of change in a simulation instantiating two experimentally-supported properties of language:
 1. Pathways of change in usage are influenced by the existing set of categories.
 2. Abstraction of category labels is influenced by the input distribution -- which evolves through usage in the previous generations.