The Roles of Acquisition and Usage in Morphological Change*

CLAY BECKNER and ANDREW WEDEL
University of New Mexico and University of Arizona

0. Introduction
There is a century-old question in linguistics concerning the role child language acquisition may play in language change (Jespersen 1922). Many have argued that language changes exclusively through the accumulation of errors in acquisition over many generations. Henry Sweet expressed this view in 1899: ‘If languages were learnt perfectly by the children of each generation, then languages would not change...changes in language are simply slight mistakes, which in the course of generations completely alter the character of the language’ (cited in Jespersen 1922:161). However, others have just as forcefully argued that ‘babies and very young children are fairly irrelevant’ to language change (Aitchison 2001:210), pointing to evidence that change occurs gradually among adolescents and adults.

The present paper revisits the relationship between acquisition and change, using computer modeling as an investigative tool. We focus in this study on which mechanism(s) of change best account for the well-known tendency in morphology for frequent forms to be suppletive or otherwise irregular (Corbett et al. 2001). For our present purposes, we are interested in an opposition between two hypothesized mechanisms of change, which we briefly review below: language acquisition, and language usage.

Language acquisition has assumed a central role in generative accounts of language change since the 1960s (Halle 1962). Generativists argue that acquisition must be the source of language change, because adults can make only superficial changes to their language, and cannot alter their grammar (Lightfoot 2006). This outlook is tied to evidence regarding the ‘critical period,’ indicating that an individual’s ability to learn a language fluently declines sharply beyond childhood (Lenneberg 1967). Moreover, it is held that errors during acquisition would be unsurprising because of the ‘poverty of the stimulus’: children must deduce I-language (a mental grammar, with infinite capacity) on the basis of noisy and incomplete data from E-language (a finite set of actual utterances) (Chomsky 1980). Outside the domain of generative theory, the Iterated Learning Model

* The order of authors is alphabetical.
Clay Beckner and Andrew Wedel

(ILM) (Smith, Kirby and Brighton 2003) has taken a different stance, while still giving acquisition a central role in language change. While leaving intact the premises of the poverty of the stimulus argument, the Iterated Learning literature argues against the nativist elements of Chomskian theory, holding instead that language adapts to be learnable in the face of a ‘transmission bottleneck’: the limited number of E-language utterances from which children must acquire an open-ended I-language capacity. ILM theorists argue that certain features of language, such as compositionality and regularity, are in fact adaptations by language to the fact that it is culturally transmitted across generations (Smith et al. 2003).

In contrast, language usage as a mechanism of change involves continuous, gradual adjustments to language structure across the lifespan of each individual. In usage-based theory (Langacker 1987, Bybee 2006), there is an ongoing, bidirectional influence between linguistic utterances and linguistic representations, and grammatical competence is thus not isolated from knowledge of linguistic usage. Rather than focusing only on changes in abstract grammar, usage-based approaches take an inclusive approach to change: any change in usage (including frequency of some variant) corresponds to a change in linguistic representation. Usage-based theorists acknowledge that language-learning abilities decline with age, but argue that this does not mean that adult grammar is ‘frozen’ (Croft 2000:58). Proponents of usage-based change point to evidence that adult speakers adopt ongoing changes in their language (Harrington 2006), that adults have innovated systematic grammatical conventions in L2 speech communities (Sankoff and Laberge 1973), and that adults most likely originate grammatical features that are acquired late by children (Bybee 2009).

For illustrative purposes, we have here presented language acquisition and language usage as qualitatively distinct, mutually exclusive mechanisms of change. However, the real-world situation is undoubtedly more subtle. There is no reason to assume that language change is driven exclusively via acquisition, or exclusively via usage. Indeed, proponents of the ILM indicate that the model may be extended to incorporate intra-generational change (Smith et al. 2003:374). Moreover, usage-based theory leaves open the possibility that acquisition plays a small role in some types of language change (Bybee 2009). Nevertheless, it is worthwhile to study the mechanisms of language change in detail, to consider which types of change are more likely to be brought about via acquisition, by usage, or both, and what the relative importance of each mechanism may be. The goal of our present project is thus to investigate such dynamics systematically using computer simulations. Computer simulations are a useful methodology in the present context, since they allow us to observe trajectories of change, and test hypotheses while controlling variables that are difficult or impossible to control in the real world (Cangelosi and Parisi 2001). With this approach, we may explore counterfactual thought experiments that are otherwise not possible. What patterns may arise if the only avenue permitted for change is via usage—if no new children are ever born, and adults can change their linguistic representations conser-
Acquisition and Usage in Morphological Change

vatively? Alternately, what patterns may arise if we assume the only avenue for change is via acquisition—if no changes in representation are permitted beyond an acquisition stage?

In the present paper, we investigate these alternatives with respect to the relationship between frequency and morphological irregularity. It is well known that over time, infrequent items are more likely to undergo analogical leveling in morphology (Paul 1890, Hooper 1976, Lieberman et al. 2007), with the result that irregular morphology tends to be concentrated among the highest-frequency items of a language (Corbett et al. 2001). What causes this pattern? Some views focus on intergenerational transmission, since rare irregular items may never be encountered during acquisition, and are thus regularized to fit a more general pattern. Thus, in discussing an ILM simulation by Kirby (2001), Dowman, Kirby and Griffiths (2006:89) write that the distribution of irregular forms ‘can be explained in terms of adaptation to cultural transmission. Put simply, frequent verbs can afford to be irregular, since they will have ample opportunity to be transmitted faithfully through the [acquisition] bottleneck.’ In contrast, usage-based approaches focus on the varying strengths of different items in memory according to their frequency. Thus, Haspelmath (2008) writes: ‘The reason why high absolute frequency favours suppletion (and irregularity more generally) has long been known: High frequency elements are easy to store and retrieve from memory, so there is little need for regularity.’ (See also Hooper 1976.)

In Section 1 of this paper, we discuss the ILM simulation approach of Kirby (2001), which shows that in principle, an acquisition-driven mechanism based in sampling-error could account for the correlation between frequency and irregularity. In Section 2, we expand Kirby’s (2001) model with an exemplar-based simulation that permits variation and gradual morphological change. We present two versions of this simulation, one in which change proceeds via intergenerational acquisition, and one in which change is usage-based. Our results indicate that either mechanism operating alone could, in principle, lead to the expected relationship between frequency and irregularity. In Section 3, we discuss additional factors that are relevant to the evaluation of our results.

1. Irregularity in the Iterated Learning Model (Kirby 2001)
The Iterated Learning Model views language change as the product of cultural transmission, in which information about a language is passed iteratively from generation to generation, with the possibility for error or innovation whenever transmission occurs (Smith et al. 2003). The ILM adopts the generative view that the information available to a language learner is impoverished (Chomsky 1980), creating a transmission bottleneck that increases the likelihood of change. This bottleneck occurs because a language learner must construct a grammar and lexicon on the basis of an incomplete sample of the language in her environment, which may by chance not include some forms. When a transmission failure occurs, information is lost about a particular item in the lexicon which then may be reconstructed on the basis of more global information available in the lexicon.
In the ILM, sampling error in acquisition accounts for the regularization of irregular morphological forms (Kirby 2001). Over many iterations of intergenerational acquisition, this model predicts that originally irregular forms will regularize probabilistically in relation to how likely they are to fail to be transmitted across generations. Assuming that irregularity is constantly introduced through other processes of change, languages will evolve toward an equilibrium in which irregulars tend to be frequent. Kirby (2001) illustrates this process using a simulation of the transfer of morphological patterns in an iterated learning paradigm. The simulation tracks the transmission of a set of bimorphemic, initially non-compositional words over the course of many generations. When a speaker needs to produce a meaning for which it has not learned a word, it reconstructs a word on the basis of any form-meaning regularities that exist among the words it has learned. In this way, any regularities that exist between words sharing meanings will spread to other words sharing those meanings at a rate proportional to their probability of loss in each iteration.

Kirby’s (2001) simulation results do indeed produce the expected results: over time, the least frequent forms regularize more rapidly, leaving irregularity to the most frequent forms. However, the results in Kirby (2001) do not allow us to draw any conclusions about the potential for usage-based change, because imperfect learning through failure of transmission is the only source of change permitted in that model. In Section 2, we describe simulations that address both intergenerational and intra-generational change in a simple exemplar-based model. We find that in a model solely allowing usage-based change, recency effects in activation levels coupled with a capacity to extend existing patterns can also cause irregularity to become concentrated among higher-frequency items.

2. Exemplar Models of Vertical and Horizontal Change

In this section, we outline two new simulations which expand on Kirby’s (2001) model in several ways, both using an exemplar architecture. The exemplar framework is based on evidence that across domains (including phonology, morphology, and syntax), linguistic representations are rich and detailed, and include memories of specific instances of use (Goldinger 1996, Bybee 2006). In an exemplar model, every token of experience has some impact on stored representations (often simply involving the strengthening of some category in memory), and the contributions of a particular token depend upon its recency, overall frequency, and similarity to other remembered tokens (reviewed in Pierrehumbert 2001). The shifting of perception and production behavior with experience can be modeled through storage of new exemplars of a category through experience, coupled with the slow decay of older exemplar memories.

An exemplar approach has the advantage of representing knowledge of linguistic variation. Speakers do in fact exhibit variation in morphology along with other domains, and an individual speaker understands (and may even produce) morphological ‘doublets’ in free variation such as *dived/dove* and *slayed/slew* (Haber 1976). In fact, competing morphological items may exist side-by-side
Acquisition and Usage in Morphological Change

within a speech community for centuries (see Fehringer 2004). An exemplar approach permits each speaker’s grammar to accommodate variants for a single paradigm slot, and allows usage preferences to shift gradually over time.

Below we compare the behavior of two exemplar-based models in simulation: an acquisition-driven, ‘vertical transmission’ model similar to Kirby’s (2001) model, and an architecturally parallel usage-driven, ‘horizontal transmission’ model. A comparison of results between the two models illustrates how low production frequency can favor the establishment of regularity both in acquisition and in usage. In the acquisition-based model, a regular variant is produced when the corresponding form was not acquired. Similarly, in the usage-based model, we can think of a pattern-extension in a production event as occurring when the lexical entry in a paradigm slot fails to supply the learned form. In this way, each production event represents a momentary bottleneck through which only one variant can pass. Error in this process creates the potential for change in a way that is conceptually similar to change via an acquisitional bottleneck.

We first describe the acquisition-driven version of the simulation. Agents possess a semantic space which is divided into 100 word categories, where each category consists of exemplars initially coded as ‘irregular.’ (Because we are abstracting away from any individual form-related influences on regularization probability, we can simply code each exemplar as ‘regular’ or ‘irregular’ without including any other phonological information.) Each word category is associated with a probability of production given by its rank order in a Zipfian distribution; if a chosen word-category contains multiple exemplars, one is randomly chosen for production. The set of exemplars in a lexical category represents the set of patterns that an agent was exposed to during acquisition, where the frequency of a particular exemplar-type in a lexical category corresponds directly to the frequency with which it was heard. In the first round, the teacher-agent transmits 500 random samples of her lexicon to a naïve learner, who stores these samples as exemplars in her corresponding lexical categories. Because sampling is random, it is possible that any word category will fail to be represented in the input to the learner in any given round. This failure is much more likely for word-categories at the low end of the Zipfian probability distribution. At the end of one round, the learner becomes a new teacher, and transmits a random 500-item sample of her lexicon to a new learner, and so on. If a teacher attempts to transmit a word-category for which she has no exemplar, she transmits an exemplar coded as ‘regular,’ representing extension of a pattern to fill in the empty paradigm slot. Once a word category contains only exemplars coded as ‘regular,’ it will remain fully regular for the remainder of the simulation, because whether this category is acquired by a learner or not, only regular forms will be transmitted in subsequent cycles. In the limit then, all word categories will become regular, but lower frequency categories decay to regularity at a faster rate. Figure 1 illustrates this with the averaged result of 100 runs of the same simulation, each iterating for 100 transmission-acquisition cycles.
The usage-driven version of the simulation is nearly identical to the acquisition-driven simulation. When run in the usage-driven mode, a simulation begins with two agents with the same 100 word categories with the same Zipfian production-probability distribution, all initialized with exemplars coded as ‘irregular.’ Each exemplar starts with an activation of 1, which is decremented exponentially in each round by multiplying by 0.5. In a given round, the speaker-agent randomly chooses a lexical entry to produce in proportion to the Zipfian production probability, and then randomly chooses an exemplar from that category in proportion to activation. The listener-agent stores that form as a new exemplar in the corresponding lexical category. After 100 productions, the roles reverse. Within this usage-based model, there are two distinct pathways for regularity to emerge from irregularity. The more intuitively obvious path to regularity lies in unequal rates of regularization error in production. Given that infrequent forms are more likely to be errorfully regularized in production (Bybee and Slobin 1982) and that usage experience leaves traces in memory that may feed back to future usage behavior, usage should contribute to regularization of infrequent forms. To illustrate this, we set the regularization probability per production event to be inversely proportional to production frequency within a regularization probability range of zero to 0.1. We ran 1000 cycles with 100 tokens exchanged in each production/perception round. At the end of 1000 cycles, a word category was counted as ‘regular’ if greater than 99% of the activation within the category derived from regular exemplars. Not surprisingly, because regularization error is greater for low frequency items, the rate of regularization for word categories under these conditions is inversely related to frequency. Figure 1 shows the results obtained by averaging over 100 such simulation runs.

The second mechanism is directly parallel to the sampling-error based mechanism of regularization in the acquisition model. All stored exemplars within a category, whether coded as ‘irregular’ or ‘regular,’ are decremented in activation identically in each round. Once a word category contains both regular and irregular exemplars, regular forms can be produced independently of error because production proceeds by choosing one of the set of extant exemplars within a category. For infrequent categories, more decay occurs between usage events, with the result that production behavior is relatively dominated by a smaller number of recent exemplars. This smaller pool of active exemplars for production makes it more likely that all irregular exemplars will be lost to decay simply through sample error in choice of templates for production. Within this simulation architecture, the contribution of sample error to regularization can be modulated by changing the exemplar decay rate – in the limit of no decay, the rate of accumulation of regular exemplars in a lexical category is simply the rate of errorfully regular production. If exemplar decay is included, as in the results shown here, the

---

1 To our knowledge, there is no direct experimental evidence that use of particular morphological variants are primed by recent experience. However, there is evidence that recent experience primes future use of both phonetic (Goldinger 2000) and syntactic (Loebell and Bock 2003) variants, so it is plausible that the same effect will hold for morphological variants.
Acquisition and Usage in Morphological Change

average rate of accumulation of regular exemplars in lexical categories is considerably faster than this. The two mechanisms together represent two intuitions about the effect of low frequency: less experience with an irregular form both makes regularization error more likely, and also increases the relative influence of recent experience on present behavior. Conversely, highly frequent irregular patterns are less likely to be regularized in production, and should also be less susceptible to destabilization by recently perceived regularization errors.

(1) Category regularity by production frequency averaged over 100 independent simulation runs.

The reduced rate of regularization seen in Figure 1 for the least frequent forms arises because within our usage-only simulation frequency is a limiting factor in the rate of category regularization: lexical items only have the opportunity to regularize when they are used, and thus extremely low-frequency items have sparse opportunities to undergo change during a simulation run. This is the converse of the situation within the transmission-only model where change is driven through lack of usage. However, as can be seen in Figure 1, both these acquisition- and usage-based models can produce regularization in relation to production frequency. In the acquisition-driven model, regularization is favored by low frequency because low-frequency irregular forms are by chance less likely to be transmitted. In the usage-driven model, regularization is favored by low frequency both because low-frequency items are more likely to be errorfully regularized in each production event, and because the set of low-frequency

---

2 We note further that in the data presented in Figure 1, we chose a very conservative regularization rate that never exceeds 0.1; this means that even for the rarest of items in the language, speakers still accurately retrieve the irregular form 90% of the time. A more pronounced error rate for very rare items would accelerate usage-driven change in this range.
irregular exemplars in a lexical entry are more likely to be lost to memory decay. The fact that both versions produced a similar regularity-irregularity boundary in the figure above is not theoretically significant because the position of the boundary depends on a number of free parameters in the models. Our intended point is that both acquisition- and usage-driven change can potentially produce the same relationship between irregularity and frequency.

3. Discussion
The simulations described in Section 2 were intended to investigate mechanisms of change that may cause irregular forms to be overrepresented among frequent items. We find that either acquisition or usage is a conceptually viable mechanism for producing the pattern in question. In our first simulation, we replicate the result from Kirby (2001), with the enhancement of an exemplar architecture that incorporates variation. However, our second simulation investigates usage-based change which proceeds in the absence of acquisition, and arrives at a qualitatively similar result. The results from this model thus constitute a computational ‘existence proof’ consistent with a usage-based mechanism (paralleling a corresponding accomplishment in Kirby (2001) for an acquisition-driven approach). These results confirm that either usage or acquisition is a theoretically possible mechanism of change behind the distribution of irregular forms.

However, a full evaluation of these mechanisms requires that we examine additional factors, including the plausibility of certain assumptions implemented within the computer simulations. We first note that a crucial assumption in the acquisition-driven version of our simulation, as well as the ILM simulation of Kirby (2001), is that morphological acquisition is based on sparse data. Morphological change arises in these simulations only because of a transmission bottleneck: due to sampling error, particular forms may never be encountered during acquisition, requiring the learner to create those forms by reference to a general pattern. There is a long history of characterizing language acquisition input as impoverished, although such views have in fact become controversial (e.g., see Pullum and Scholz 2002). Morphology may be a particularly questionable domain in which to argue that acquisition data is highly impoverished. For morphological input, there is no ‘poverty of the stimulus’ in the usual Chomskian sense, because morphology does not represent the same kind of seemingly unbounded system as syntax (Spencer 2001:126).

It is indeed true that during acquisition, children do not encounter every combinatoric morphological possibility, especially in the case of languages having complex systems of inflectional or derivational morphology (Albright 2008). However, how likely is it that such acquisitional deficits account for all regulari-

3 For the acquisition-driven model, these parameters are the sample-size in each transmission event and the number of such events. For the usage-driven model, the position of the boundary is influenced by the frequency of switching speaker/hearer roles, the rate of exemplar-decay, the probability of regularization in each production event, and the number of such events, i.e., usage frequency. In both versions, these factors influence the rate at which error can modify a category.
Acquisition and Usage in Morphological Change

zation events over the course of language change? Even in morphologically simple languages (such as English), we see many high-frequency regular forms (e.g., past tense forms such as wanted, asked, looked, played). For such forms, it is implausible they could remain unlearned a sufficient proportion of the time to drive language change. We note that in order for our present acquisition-driven simulation to produce reasonable results, it must be set up with an extremely ‘tight’ transmission bottleneck. At each generation, the training set (500 tokens) is only five times the total lexicon size (100 items to express, with a Zipfian distribution), and a similar bottleneck is applied in Kirby (2001) (50 training items for 16 morphological combinations). Such tight bottlenecks clearly exaggerate the nature of data constraints on L1 acquisition. Conversely, if simulations are run with weaker bottleneck, i.e., with a high probability of exposure to medium-frequency forms in acquisition, irregularity remains the norm in all but the lowest-frequency forms.

A further point to consider is that in order for acquisition to have an effect on language change, the innovations of young children would need to persist in the population. However, when these matters have been investigated for morphology, the findings pose problems for the acquisition-driven view.4 Bybee and Slobin (1982) studied the morphological errors of three age groups: preschool children, 8-10 year-olds, and adults. There are similarities across all the groups, since all speakers tend to regularize infrequent forms, and speakers of all ages may thus contribute to this process of historical change. However, only older children and adults produce the past tense in a way that is consistent with the ongoing pattern of change in English, namely, the regularization of -t and -d final verbs (hit > hitted). In contrast, Bybee and Slobin (1982) find that preschool children produce zero-marking errors on this same class of verbs, which is a fleeting innovation with no reflection in the current direction of change in English.

Why might it be that innovations by young children fail to take hold in a speech community? As a general rule, language changes only persist when they originate in socially influential groups, yet small children do not constitute such a group (Kerswill 1996). A lengthy acquisition process allows young learners to conform to the norms of their speech community by first converging on the speech of their parents. Sociolinguistic research finds that in-progress language change tends to be most prominent in speakers in late adolescence—not among young children as we would expect if imperfect learning were the primary mechanism of change (Tagliamonte and D’Arcy 2009). The propagation of innovation thus presents difficulties for accounts of change that are solely based on acquisition (Croft 2000). Social propagation of innovations is an important area for future enhancements of the present simulations, since as a simplification in the current model (following Kirby 2001), interactions occur only between idealized pairs of speakers rather than in a speech community.

4 For reviews of related evidence in other domains of language, see Aitchison (2001), Croft (2000).
The foregoing concerns present significant challenges for any model that attributes morphological change principally to imperfect learning in acquisition. Instead, we suggest that distinct mechanisms of change in acquisition and usage may independently contribute to the observed concentration of irregularity among high-frequency items. As we have argued, it is unlikely that a failure to transmit items intergenerationally accounts for all instances of regularization. However, such failures clearly do occur with respect to the rarest items in a language. Critics of the imperfect learning model acknowledge that acquisition accelerates usage-driven change in this way, by finalizing the loss of variants that are already extremely rare (Aitchison 2003:739, Bybee 2009:349). As a consequence, it may be that analogical extension in usage may be primarily responsible for regularization in the mid- to high-frequency range, while failure in transmission may become an important source of regularization in the lowest frequency range. Future simulation research in this area will investigate more fully the interaction of usage and acquisition in morphological change.

With respect to analogical change in morphology, we have described an inclusive account in which usage is a central factor, and in which acquisition likely plays a supportive role. Yet care must be taken to avoid generalizing the details of these dynamics to all domains of language change. There is no requirement that usage and acquisition influence language in the same direction, and indeed their influences may often compete. It is sometimes claimed that language adapts to be learnable (Smith et al. 2003:385), but we note that certain learnability factors may be diametrically opposed to usage factors. For instance, phonetic reduction is most prominent in highly frequent units as a result of extensive articulatory practice. The resultant reduced forms may in fact be harder to learn, because they are articulatorily more complex and morphologically more opaque (see Bybee 2009). A full account of language change will thus acknowledge that the interactions between usage and acquisition are quite complex, and under different circumstances the contributions of these mechanisms may amplify or compete with one another.

References

Acquisition and Usage in Morphological Change


Clay Beckner
University of New Mexico
Department of Linguistics
MSC 03 2130
Albuquerque, NM 87131
coogle@unm.edu

Andrew Wedel
University of Arizona
Department of Linguistics
Tucson, AZ 85721-0028
wedel@email.arizona.edu