

This poster examines loanword adaptations from a Bayesian perspective, and argues that stochastic behavior is located in the learning process. Iverson and Lee 2004 survey English words borrowed into Korean, and show that final [t] is adapted in several ways, including /t<sup>h</sup>/ and /s/, (1). (Final /s/ is justified by the object forms.)

- (1) a. adapted as /t<sup>h</sup>/ 'cut' ==> [k<sup>h</sup>ət<sup>h</sup>i]  
           'net' ==> [nɛt<sup>h</sup>i]  
       b. adapted as /s/ 'internet' ==> [int<sup>h</sup>ənɛt], [int<sup>h</sup>ənɛsil] object  
           'Hamlet' ==> [hɛmnit], [hɛmnisil] object

Korean morphemes can end in /t/, e.g. /nat/ 'cereal grains' (Sohn 1999: 165) but very few do. We seek to connect this statistical fact to (1). Two principle-based approaches in OT make the wrong predictions. Lexicon Optimization ensures that learners "will always choose lexical forms that can receive faithful parses given their language's syllable inventory" (Prince and Smolensky 2004: 227). This is violated in (1) when final /s/ is chosen, and the more faithful /t/ rejected, for the English models never contain [s]. The Free Ride Principle (McCarthy 2004: 1) is a possible replacement for Lexicon Optimization: "learners who have discovered the /A/ -> [B] unfaithful map from alternations will attempt to generalize it, projecting /A/ inputs for surface [B]s, whether they alternate or not." But Korean offers too many free rides: /t<sup>h</sup> s c<sup>h</sup> h/ all neutralize to [t] word-finally, (2) (Sohn 1999: 165), so how will the learner choose?

- (2) a. [nat] < /nat<sup>h</sup>/ 'piece, unit'  
           < /nas/ 'sickle'  
           < /nac/ 'daytime'  
           < /nac<sup>h</sup>/ 'face'  
       b. [hiit] < /hi-ih/ the name for the letter ㅎ

We offer instead a Bayesian analysis-by-synthesis account. Bayes's rule is derived in (3); its application to McCarthy's schematic example is in (4).

- (3)  $p(A | B) p(B) = p(A \& B) = p(B | A) p(A)$  therefore  $p(A | B) = p(B | A) p(A) / p(B)$   
 (4)  $p(/A/ | [B]) = p([B] | /A/) p(/A/) / p([B])$

In (4) the probability of choosing underlying /A/ given the learner heard [B] is proportional to the probability that they would pronounce [B] from /A/ (this is the analysis-by-synthesis part) and the probability of /A/, either derived from statistics over the existing lexicon, or in the case of the earliest word-learning from general markedness principles for underlying representations (a Bayesian prior). For the alternatives, /A<sub>1</sub>/, /A<sub>2</sub>/, ... we will calculate odds ratios across the hypothesis space. In (5) we calculate the odds ratio for /s/ versus /t/. The term p([t]) is the same in both cases and so cancels; likewise in final position  $p([t] | /s/) = p([t] | /t/) = 1$  as both /t/ and /s/ obligatorily neutralize to [t].

- (5)  $p(/s/ | [t]) / p(/t/ | [t]) = [p([t] | /s/) p(/s/) / p([t])] / [p([t] | /t/) p(/t/) / p([t])]$   
            $= p(/s/) / p(/t/)$

Bayesian learning thus says that the odds of choosing an underlying form should favor the most probable form given the existing lexicon, but that learners can sometimes "bet on" other less probable forms, as indeed is exhibited in the loanword adaptations. Given this Bayesian account, principles like Lexicon Optimization and Free Rides are rendered superfluous.