

**The role of phonetic knowledge in phonological patterning  
corpus and survey evidence from Tagalog infixation**

Kie Zuraw

University of California, Los Angeles

UCLA Department of Linguistics

3125 Campbell Hall, Box 951543

405 Hilgard Ave.

Los Angeles, CA 90095-1543

[kie@ucla.edu](mailto:kie@ucla.edu)

**The role of phonetic knowledge in phonological patterning**

**Corpus and survey evidence from Tagalog infixation**

## ABSTRACT

A current controversy in phonological theory concerns the explanation of crosslinguistic tendencies. It is often assumed that crosslinguistic tendencies are explained by mental bias: a pattern is common because it is favored by learners/speakers. But work by Blevins and colleagues in *Evolutionary Phonology* has argued that many crosslinguistic tendencies can be explained without positing such bias. This would mean that crosslinguistic tendencies cannot be unproblematically used as evidence about the mental machinery that humans bring to learning and using language. In response, many researchers have looked at different types of data, such as processing, learning of real and artificial languages, and literary invention. This paper presents another type of data: extension of native-language phonology to words with novel phonological structure, in this case infixation in Tagalog into loanwords with novel initial consonant clusters. The data come from a written corpus and a survey. Tagalog speakers' treatment of these clusters parallels crosslinguistic findings of cluster splittability by Fleischhacker. This paper argues that explaining the data requires attributing to Tagalog speakers phonetic knowledge and a bias about how to apply that knowledge.\*

\* Many people besides the author have put substantial work into this paper. For detailed critiques of every aspect of the paper, I'm indebted to Brian Joseph, Jaye Padgett, Donca Steriade, and two anonymous reviewers. This project was prompted by and draws heavily on the research of Heidi Fleischhacker MacBride. Essential to the creation of the corpus used in this paper were programming work by Ivan Tam; a grant from the UCLA Faculty Senate; and earlier work by Rayid Ghani, Rosie Jones, and Dunja Mladenec, who generously shared their corpus. For

discussions and suggestions on various components of the project, I thank Adam Albright, Edward Flemming, Bruce Hayes, Angelo Mercado, Christian Uffmann, Shelley Velleman, Colin Wilson, and audiences at the GDR Phonologie, NYU Workshop on Redefining Elicitation, LSA, OCP II, AFLA XII, UC Berkeley, UC Santa Cruz, MIT, and University College London.

**The role of phonetic knowledge in phonological patterning  
corpus and survey evidence from Tagalog infixation**

Kie Zuraw

**1. Introduction.** Generative linguistics seeks to describe the mental apparatus (language-specific and otherwise<sup>1</sup>) that humans bring to the task of learning and using language. In the realm of phonology, at least, this inquiry most often takes the form of asking whether learners favor some conceivable grammars over others. The challenge lies in determining which pieces of evidence actually bear on the question of learner preferences, and which are to be explained by other means. To take a simple example that has been discussed before (see Steriade 2001a, Hura et al. 1992), many languages assimilate a nasal consonant's place to that of a following obstruent (/an+pa/ → [ampa]), but not a preceding obstruent (/ap+na/ → [apna]). This typological observation is accompanied by a functional observation, in this case a phonetic one: a nasal's place of articulation is more difficult to perceive in the environment *vowel\_\_obstruent* than in the environment *obstruent\_\_vowel* (for most places of articulation). But how does the phonetic observation translate into an explanation for the typology?

One possibility is that humans' cognitive apparatus encodes the undesirability of maintaining place where it is hard to perceive. First, we must be able to learn in what environments nasal place is hard to perceive (or perhaps be endowed innately with this knowledge). And second, we must be biased against maintaining hard-to-perceive place contrasts. Under this approach, the functional motivation—phonetic knowledge plus a bias about how to apply it—is inside the mind. This is the position taken explicitly by Steriade 2001a, for example, and is implicit in many other works (see Hayes & Steriade 2004). More generally, the

idea that typological tendencies are to be explained by mental biases has pervaded generative phonology at least since Chomsky and Halle 1968.

A second possibility, however, involves language transmission: because nasal place is hard to perceive in the *vowel\_\_obstruent* environment, learners will have a tendency to mis-hear /an+pa/ as [ampa],<sup>2</sup> but to correctly hear /an+i/ as [ani]—that is, to mis-hear the morpheme /an/ as alternating between [am] and [an]. If this misperception is widespread enough, it will appear to learners that the language has a process of nasal place assimilation to a following obstruent, and this will be encoded in the learner’s grammar. Thus, languages without assimilation can change into languages with assimilation, and the change will be more frequent for preobstruent assimilation than for postobstruent assimilation, since misperception is less likely in the *obstruent\_\_vowel* environment. Under this approach, the functional motivation for the typological trend is outside the mind. Humans need not have any knowledge of perceptibility, let alone a bias about how to apply that knowledge. This is the position advanced by Blevins and Garrett (1998, 2004), Blevins (2004) within the framework of Evolutionary Phonology. See also Ohala 1981, 1993, and others; Hale and Reiss 2000; Hyman 2001; Myers 2002; and Yu 2003, 2004.

Work in Evolutionary Phonology and in the same spirit has included two strands: explanations for functionally motivated ‘natural’ typological patterns that seemingly remove the need for positing phonetic knowledge or bias (e.g. the work by Ohala); and examples of ‘unnatural’ patterns (along with diachronic explanations of them) to show that they must also be learnable (e.g. Hyman 2001, Yu 2004). For example, standing against the many languages with postnasal voicing of obstruents (see Pater 1999; see Hayes & Stivers 1995, Hayes 1999 for an aerodynamic motivation), Hyman gives a case of postnasal devoicing of obstruents (though

Zsiga, Gouskova, & Tlale 2006 argue that the language in question, Tswana, does not phonetically have postnasal devoicing: of the six speakers they recorded, some have devoicing of stops across the board, some have no devoicing at all, and some devoice everywhere but word-initially).

The existence of these unnatural cases—if the ‘unnatural’ analysis is the correct one—is important, because it rules out certain hard-line positions.<sup>3</sup> For example, under the classic Optimality Theory (OT) idea that the constraint set is universal (Prince & Smolensky 1993/2004), we might want to say that only functionally motivated constraints belong to that set, and thus only ‘natural’ languages are possible. If ‘unnatural’ languages do exist, this position is not tenable, and if the language faculty does include substantive biases, at least some of them must be only that—biases—and do not rule out as unlearnable all contrary languages. See Wilson (2006) for a development and implementation of the idea of soft biases within a constraint-based framework.

The strand of the Evolutionary Phonology program that seeks to explain typological trends has shown that it is dangerous to make inferences about substantive biases from typology, because typological patterns may result not from those biases but from tendencies in language transmission. One response to this situation is to continue to investigate, in individual cases, whether an account of a typological tendency is constructible without implicit knowledge or bias; another is to test hypotheses about mental biases using other types of data.

Many researchers, in seeking other types of data, have probed speakers’ behavior in situations where it is not directly determined by their native-language experience, so that the history that shapes that experience cannot be an explanation for the behavior (another is to probe processing of ‘natural’ vs. ‘unnatural’ native-language phonology, as in Zhang & Lai

(submitted), Zhang, Lai & Turnbull-Sailor (in progress)). This type of research has included artificial language-learning experiments (Guest, Dell & Cole 2000; Pater & Tessier 2003; Pycha & al. 2003; Wilson 2003, 2006), including novel language games (Treiman 1983, Derwing & al. 1988, Pierrehumbert & Nair 1995), and the study of second-language phonology (Broselow 1992a, 1992b). Less commonly, there has been research on literary invention, such as puns, rhymes, and alliteration, mostly using corpora (Minkova 2001, 2003; Fleischhacker 2002b, 2005; Steriade 2003; Kawahara to appear). The study of the phonological adaptation of loans also falls into this category, though interpreting the data is made more difficult by the question of what borrowers perceive (e.g. Silverman 1992, Yip 1993, Dupoux & al. 1999), and uncertainty as to the mechanism of borrowing (directly from foreign speakers or mediated by bilinguals), the degree of contact at the time of borrowing, the social context of the borrowing, etc. Least commonly, there has been research on the extension of authentic native-language grammar to unprecedented cases—that is, not just the application of native-language grammar to novel words (the wug-testing pioneered by Berko 1958), but its application to novel types of words. The English plural-of-Bach test proposed by Lise Menn (Halle 1978) would be an example: is it [baxz], [baxs], or [baxɪz]? This article aims to contribute to the debate on substantive biases in the language faculty by presenting evidence from a study of this last type, involving infixation in Tagalog stems with novel initial clusters. It is argued that the Tagalog evidence supports the existence of a mental bias.

As in most of the works just cited, the structure of the argument is along the same lines as Pullum & Scholz's (2002) definition of argument from poverty of the stimulus (see section 6).<sup>4</sup> That is, speakers are argued to have implicit knowledge that they could not have acquired, given the data available to them, unless they brought a certain prior bias to the learning task. Thus, the



existence of that prior bias is supported. The phenomenon in question is infixation into stems beginning with consonant clusters in Tagalog. The infix may split the cluster (g-um-raduate) or not (gr-um-aduate), with the frequencies of the two variants depending on the consonants in the cluster.

In what follows, I first review previous findings on cluster splittability, with an extended discussion of Fleischhacker's (2002a, 2002b, 2005) perceptual-similarity account, and explain the relevance of Tagalog infixation (section 2). I then present evidence from a written corpus of Tagalog (section 3), and from a survey of Tagalog speakers (section 4). It is argued that both the corpus and the survey evidence follow a predicted crosslinguistic pattern, that an explanation based on language transmission is unlikely, and that therefore Tagalog speakers do have phonetic knowledge of consonant clusters and a bias about how to apply that knowledge. I then sketch an OT analysis, which includes a proposal about the form of constraints that regulate similarity between related surface forms (section 5), and finally consider alternative explanations of the data (sections 6, 7, and 8).

## **2. Cluster splittability**

**2.1. Previous findings.** There is much previous research on how word-initial consonant clusters behave in situations where the cluster could become split. The most extensive evidence comes from epenthesis in loanword adaptation or second-language phonology, and the most robust finding there has been that stop-sonorant clusters (TL) are more splittable by an epenthetic vowel than are sibilant-stop clusters (ST) (Fleischhacker 2002a; Broselow 1983, 1992a, 1992b; Singh 1985). The pattern found in Farsi (from Fleischhacker 2002a; see also Karimi 1987, Shademan 2002) is typical. Foreign words beginning with an ST cluster receive an initial

prothetic vowel, leaving the cluster intact, as in esparta ‘Sparta’, whereas words beginning with a TL cluster receive an epenthetic vowel that splits the cluster (anaptyxis), as in pelutus ‘Plutus’. The pattern is repeated in many other languages, and the reverse does not seem to be attested.

To explain this anaptyxis-prothesis asymmetry, representational approaches have proposed that ST forms a structure more cohesive than TL, such as a complex segment or linked structure (Fudge 1969, Ewen 1982, Selkirk 1982, Broselow 1992b, van de Weijer 1996, and others). If this structure is illegal in the borrowing language, but also resists splitting, then ST can neither be tolerated nor be split, and prothesis occurs (ST... → VST...). Under these accounts, splitting is the norm, but ST resists it. Some representational approaches have attributed ST onsets’ special structure to their falling sonority profile, or to a shared laryngeal gesture (Broselow 1992b, following Browman and Goldstein 1986). See section 8.1 for an attempt to construct an articulatory account along different lines. Gouskova (2003) appeals to the markedness of the result of epenthesis, noting the differences in syllable contact produced by prothesis of words beginning with different cluster types. Assuming, following Venneman 1988, that coda-onset sequences should be of falling sonority (l.b, not b.l), Gouskova notes that prothesis of a ST-initial word produces the unmarked syllable contact S.T (VS.T...), but prothesis of a TL-initial word produces the marked syllable contact T.L. For Gouskova, prothesis is the norm, but TL clusters—and others of rising sonority—are forced to split.

The explanation of Fleischhacker (2002a, 2002b, 2005), which I adopt, is based on perceptual similarity. Fleischhacker proposes that borrowers of new words—that is, speakers of the borrowing language who have access to the form in the source language—attempt to keep the borrowed form perceptually similar to the source form,<sup>5</sup> and that TL and TVL are more similar to each other than are ST and SVT. The similarity claim is supported by experimental evidence,

summarized below. Fleischhacker speculates as to why TL and TVL should be more similar than ST and SVT, but testing that speculation is beyond the scope of her investigation, and it is not tackled here, either. Fleischhacker's speculation relies on the idea of the 'perceptual break' created by the onset of formant structure, as at the transition from T to L. The higher the intensity of the formant structure after the break, the stronger the break; the higher the intensity of the aperiodic noise before the break, the weaker the break. Thus, TL and ST are two extreme cases. TL begins with silence (though T's release burst precedes the break) and proceeds to the strong formant structure of L; the break between T and L is therefore strong. ST, on the other hand, has considerable aperiodic noise preceding the break (S) and proceeds to silence, with no formant structure at all, so the break between S and T is weak. Fleischhacker assumes that splitting a cluster at a stronger perceptual break creates a smaller perceptual departure from the unsplit original; therefore, TL and TVL should be perceived as more similar than ST and SVT.

The remainder of this section summarizes Fleischhacker's findings for clusters other than TL and ST and for phenomena other than loan epenthesis, and her experimental evidence on perceptual similarity.

The fact that TL and ST differ in both  $C_1$  (stop vs. sibilant) and  $C_2$  (liquid vs. stop) makes it hard to pin down the source of the difference in behavior. Examining sibilant-C clusters permits a more controlled comparison, since one can hold  $C_1$  relatively constant—in the examples below, mostly [s] with some [ʃ] and [z]—and vary  $C_2$ . This is what Fleischhacker (2002a, 2005) does, looking again at epenthesis in loan adaptation and creoles, where source languages have a variety of SC clusters. Among languages that tolerate no initial CC clusters, repairing them all by either prothesis or anaptyxis, Fleischhacker discovers an implicational

hierarchy, schematized in (1). Within a given language, if one of the clusters in (1) splits, clusters to the right of it must also split, as summarized in Table 1.

- (1)  $ST < S_m < S_n^6 < Sl < SR, SW$   
 less splittable                      more splittable  
 (S = sibilant; T = stop; R = rhotic; W = glide)

As schematized in Table 1, Wolof (Ka 1985, Broselow 1992b) differentiates ST, with prothesis, from the rest of the clusters, which show variation between prothesis and anaptyxis. A cut-off after S<sub>m</sub> is exemplified by Hindi (as described by Bharati 1994: 56-59), with prothesis for ST and S<sub>m</sub>, and variation or anaptyxis for the rest. Kazakh (Sulejmenova 1965: 76-83) has a cut-off between ST (prothesis) and S<sub>m</sub> (variation), and also between S<sub>n</sub> (variation) and Sl (anaptyxis). Farsi has its cut-off between Sl and SR (which Hindi as described by Bharati also differentiates).<sup>7</sup> For the other languages identified by Fleischhacker, the information is sparser but still consistent with (1).

INSERT Table 1 ABOUT HERE

The scale of splittability in (1) is expected given Fleischhacker's speculation about perceptual breaks: the SC<sub>2</sub> clusters further to the right in the scale have a C<sub>2</sub> with stronger formant structure, so the break between S and C<sub>2</sub> should be stronger. Or, to take a slightly different view, the more sonorous C<sub>2</sub> is, the more vowel-like it is, and thus the more the transition from S to C<sub>2</sub> is already similar to a transition from S to a vowel.

The influence of  $\underline{C}_1$  is less clear. Fleischhacker finds only three languages that show a difference between  $\underline{XC}_2$  and  $\underline{YC}_2$ , with  $\underline{C}_2$  held constant. One is Farsi, where stop-l clusters split (pelastik ‘plastic’ Shademan 2002), but sibilant-l clusters show prothesis. This suggests that  $\underline{TC}_2$  is more splittable than  $\underline{SC}_2$ . Similarly, Wolof splits all  $\underline{TL}$  and  $\underline{TR}$  clusters, but shows variation for  $\underline{SL}$  and  $\underline{SR}$ .<sup>8</sup> In Kirgiz, as discussed in detail by Gouskova (2003), there is at least one  $\underline{TC}_2$ - $\underline{SC}_2$  pair, kv (anaptyxis) vs. zv (prothesis), and in general, lower sonority of  $C_1$  correlates with greater splittability. The inventory of clusters borrowed into Kirgiz is rich, and Gouskova finds that clusters with falling or level sonority undergo prothesis, but those with rising sonority undergo anaptyxis, avoiding the bad syllable contact that would arise from prothesis. The falling- and level-sonority clusters borrowed into Kirgiz include not just  $\underline{ST}$ , but also rt, lb, lv, zv, and mn—all undergo prothesis. The rising-sonority clusters include not just  $\underline{Sm}$ ,  $\underline{Sn}$ ,  $\underline{Sl}$ ,  $\underline{SR}$ , and  $\underline{TL}$  (stop-sonorant), but also kv, mr, kn, and pn—all undergo anaptyxis. To differentiate the predictions of the syllable-contact account from those of the perceptual-break account, we would need data from a language with a rich cluster inventory and some cluster-splitting phenomenon that does not create a heterosyllabic  $\underline{C}_1.C_2$  sequence, such as  $C_2$  deletion.

The Tagalog data to be discussed in this paper bear only on  $\underline{SC}$  and  $\underline{TC}$  clusters. We can incorporate the Farsi and Wolof facts into the splittability scale by adding a second dimension, as in (2).

$$(2) \quad ST < Sm < Sn < Sl < SR, SW$$

^     ^

Tl    TR

Fleischhacker (2002b, 2005) presents additional evidence for a TL vs. ST difference from reduplication, imperfect puns, and alliteration. The reduplication evidence comes from languages that do not always copy a complex onset in full (i.e. ba-bladupi; see Steriade 1988 for a survey). Fleischhacker's focus is on languages with a 'restricted skipping' pattern, where some but not all clusters undergo simplification. All of the surveyed languages with restricted skipping simplify only obstruent-sonorant clusters (ba-bladupi); other clusters are either copied in full (sta-stalumi) or not copied at all (*\_e-stalumi*). Gothic, for example (Wright 1910/1954: 147-148, see also Cairns and Feinstein 1982, Broselow 1992b), copies only the first consonant of most clusters, as in faí-fráís 'tempt-preterite', gaí-grōt 'weep-preterite', and saí-slēp 'sleep-preterite'. The clusters st and sk, however, are copied in their entirety, as in ga-staí-stald 'possess-preterite' and skaí-skáip 'sever-preterite'.<sup>9</sup> Fleischhacker assumes that, as with loan adaptation, there is a preference to keep two forms similar, here the reduplicant and its base. Fleischhacker's view of Gothic and similar cases is that the pairs TV-TLV and SV-SLV are treated by speakers as sufficiently similar to allow simplification in reduplication, but SV-STV is not.

Fleischhacker draws further evidence for parts of (2) from a corpus<sup>10</sup> of English imperfect puns—puns juxtaposing two forms that are not perfect homonyms. Puns like blown apart ~ Bonaparte,<sup>11</sup> where a stop-liquid-vowel sequence and a stop-vowel sequence are compared, are more frequent than expected. That is, among puns in the corpus of the form  $C_1C_2V... \sim C_1V...$ , 40% are of the blown apart ~ Bonaparte type, with  $C_1$  a stop and  $C_2$  a liquid (TL)—whereas among all English word pairs of the form  $C_1C_2V... \sim C_1V...$ , only 26% are of that type. Pairs like sturgeon ~ surgeon, where  $C_1$  is a sibilant and  $C_2$  a stop (ST), are by contrast underrepresented, and SL Pairs like slalom ~ solemn, are somewhere in between (somewhat underrepresented). This supports the SL>TL and ST>SL comparisons in (2). (Fleischhacker does

not further break down the sonorants, e.g. into nasals, liquids, and glides.) Looking at puns of the form  $C_1C_2V... \sim C_2V...$ , on the other hand, such as Stabitha ~ Tabitha, Fleischhacker finds that they occur about as often as expected for TL, SL, and ST. The lack of distinctions found among  $C_1C_2V... \sim C_2V...$  puns is contrary to what would be expected if a structural account applied: if ST clusters were less splittable for a structural reason, they should be so regardless of whether they are split by deleting the first consonant or the second.<sup>12</sup>

Fleischhacker also gives evidence for parts of (2) from poetic alliteration, following Kuryłowicz 1971 and Broselow 1992b. In early Germanic (for which Fleischhacker cites Kuryłowicz), ST clusters alliterate only cohesively—that is, with themselves; for example, a word beginning in st, such as Old English stān ‘stone’ can alliterate only with words beginning in st, not with words beginning in sV or, say, sp. Words beginning in other  $C_1C_2$  clusters, however, alliterate with any word beginning in  $C_1$ : brim ‘sea’ alliterates with beorgas ‘hills’ and blīcan ‘shine’. Assuming that successful alliteration requires similarity between the corresponding onsets, the Germanic pattern supports the distinction in (2) between ST and the rest. Early Irish (for which Fleischhacker cites Murphy 1961) is the same as early Germanic, except that sm also can alliterate only cohesively. This provides a second piece of evidence, alongside epenthesis in Hindi, for a distinction between Sm and Sn. In Middle English, discussed in detail from a perceptual-similarity perspective by Minkova (2001, 2003), a word beginning in ST is allowed to alliterate with any word beginning in s, but it is nonetheless highly likely to alliterate cohesively. Looking at a corpus made up of three long poems, Minkova finds that 93% of st, 99% of sp, and 91% of sk alliterate cohesively. Rates of cohesive alliteration are similarly high for S-nasal—though with sn (100%) having, contrary to expectation, a higher rate of

cohesive alliteration than sm (89%)—lower for sl (68%), and lower for the remaining clusters (from 7% for fr to 50% for thr).<sup>13</sup>

Splitting in epenthesis, reduplication, the  $C_1C_2V... \sim C_1V...$  puns above, alliteration, and VC infixation has the property that if  $C_1C_2$  is split,  $C_1$  becomes vowel-adjacent ( $C_1V...$ ), as summarized in (3). Fleischhacker proposes that in all these cases, there is a preference to keep the two related forms (foreign word and loan,<sup>14</sup> base and reduplicant, etc.) perceptually similar.

(3)	unsplit		split	
	epenthesis	$C_1C_2V...$ (foreign word)	$C_1VC_2V...$	(adapted)
	reduplication	$C_1C_2V...$ (base)	$C_1V...$	(reduplicant)
	pun	$C_1C_2V...$ (one member of pun pair)	$C_1V...$	(other member)
	alliteration	$C_1C_2V...$ (one member of allit. pair)	$C_1V...$	(other member)

Assuming the scale of perceptual distance ( $\Delta$ ) shown in (4), splitting should be most likely when the difference  $\Delta(C_1C_2, C_1V)$  is small, as in (5), which restates one dimension of (2).

$$(4) \quad \Delta(C_1T, C_1V) > \Delta(C_1m, C_1V) > \Delta(C_1n, C_1V) > \Delta(C_1l, C_1V) > \Delta(C_1r, C_1V), \Delta(C_1W, C_1V)$$

$$(5) \quad \underline{\text{least splittable}} \quad CT \quad Cm \quad Cn \quad Cl \quad CR \quad CW \quad \underline{\text{most splittable}}$$

(holding C constant)

Fleischhacker's final body of evidence on similarity comes from experimental tasks. In one experiment, English-speaking subjects were asked to judge whether synthesized syllable



pairs such as klɑ ~ kɑ were the same or different; a longer reaction time or higher error rate was taken to mean that a given pair is more similar. Although Fleischhacker does not report statistical significance, one trend in the data is relevant. Among syllable pairs that delete C<sub>2</sub>, subjects discriminated T-sonorant pairs (e.g. klɑ ~ kɑ) and S-sonorant pairs (e.g. slɑ ~ sɑ) more slowly than ST (skɑ ~ sɑ) supporting the view that {S/T}LV and {S/T}V are more similar than STV and SV.<sup>15</sup> In a second experiment, Fleischhacker asked English-speaking subjects to rate the similarity of a real, CC-initial English word to a modified version with a schwa inserted. Among pairs displaying anaptyxis (like pluck ~ p[ə]luck<sup>16</sup>), subjects gave the highest ratings when the initial cluster was T-liquid, SW, or S-liquid, somewhat lower ratings for SN, and the lowest for ST (though, as Fleischhacker points out, aspiration is a confound for ST pairs such as s[ɹ]ar ~ s[ə][ɹ<sup>h</sup>]ar). This supports having T-liquid, SW, S-liquid more splittable than SN in (2), though Fleischhacker does not report whether these differences are significant. For prothesized pairs, on the other hand (pluck ~ [ə]pluck), ratings seem to be flat across the cluster types.<sup>17</sup>

To summarize Fleischhacker's findings, a group of phonological and parophonological phenomena—epenthesis, reduplication, punning, and alliteration—display a crosslinguistic trend for certain consonant clusters to be more splittable than others. There is a plausible phonetic basis for this trend, based in similarity, and some experimental support for that phonetic basis.

If the phonetic account is correct, there remains a problem in translating it into an explanation for the crosslinguistic pattern. As in the nasal-assimilation example in section 1, one possible explanation is that the phonetics are inside the mind of the speaker: speakers are able to determine how similar a C<sub>1</sub>C<sub>2</sub>-C<sub>1</sub>V pair is, and are biased to keep pairs such as foreign word and loan, base and reduplicant, etc., similar. This would follow Steriade's (2001a, 2001b) proposals concerning the P-map or perceptual map. But another possible explanation lies in language

transmission. Taking the loanword/L2 epenthesis examples, perhaps speakers are more likely to misperceive a  $C_1C_2$ -initial foreign word as having a vowel between the two Cs if  $C_2$  is more sonorous; under this account the grammar plays no role in determining where to insert vowels, and no phonetic knowledge is required of speakers.

It is less obvious how a misperception account would extend to reduplication, but perhaps learners perceive the difference between TLV... and TV... accurately in bases but inaccurately in reduplicants, for attentional or prosodic reasons. Or, the diachronic facts, if known, could suggest a nonphonetic explanation. In the case of alliteration, an account without phonetic knowledge seems less likely: Minkova's Middle English figures represent variation within the bounds of the contemporary convention, and thus appear to reflect the poets' spontaneous choices as to which word pairs produce the best alliterative effect. Likewise, Fleischhacker's pun statistics reflect case-by-case judgments by punners and their audiences as to whether a pun should be coined, and then whether it is funny enough to be repeated and thus eventually appear in an advertising slogan or a book of puns. All the pun types under consideration are possible—there is no convention that makes surgeon ~ sturgeon an illegal pun—some are merely more frequent than others.

A further finding that seems to refute a misperception or language-change account comes from an artificial language game study by Pierrehumbert and Nair 1995 (see also Fowler, Treiman & Gross 1993), in which English speakers were taught to insert VC infixes into real words. When participants were tested on words beginning with clusters, where outputs such as st-əl-ʌb or s-əl-tʌb would be possible for 'stub', and pl-ək-ænət or p-ək-lænət for 'planet', '[t]he cluster /st/ split the least, and the clusters /sl/ and /pl/ split the most.' (p. 101). The Tagalog data to be presented here—also from an infixation task—provide further evidence against an

account based purely on misperception or language change, and in favor of an account that includes phonetic bias.

Table 2 summarizes the evidence for cluster distinctions discussed in this section.

INSERT Table 2 ABOUT HERE

**2.2. Relevance of Tagalog infixation.** Certain Tagalog verbs take the infixes um and in (um is used for actor-focus forms, in for others) to mark realis aspect (um also marks infinitives), as shown in (6) (Schachter & Otones 1972, French 1988, Prince & Smolensky 1993/2004, McCarthy & Prince 1993).

(6) bago<sup>18</sup> 'new'                  **bum**ago                  'to change'

Native words in Tagalog do not have initial consonant clusters (except for some stop-glide clusters created by optional syncope; see section 8.2). Tagalog has many loans from Spanish and English that do begin with clusters, however, and these words may be infixed. Two main patterns result, as illustrated in (7): the infix may be placed inside the cluster or after it (Cena 1979, Ross 1996, Maclachlan & Donohue 1999, Orgun & Sprouse 1999).<sup>19</sup>

(7) 'graduate' **gum**graduate ~ **grum**aduate  
       'protect' **pin**rotekta-han ~ **prin**otekta-han

The situation when these loans first entered the language is similar, then, to the Pierrehumbert & Nair 1995 language game: speakers who had learned how to insert a VC infix into words beginning with a single consonant extended the pattern to words beginning with consonant clusters. This required making a decision, in each case, about whether to split the cluster. As in all the cases above, when the  $C_1C_2$  cluster is split,  $C_1$  becomes vowel-adjacent (followed by u or i). Thus, if Fleischhacker's perceptual explanation is correct, the sonority of  $C_2$  should determine the cluster's splittability.

The empirical question to be addressed here is what differences might exist in splittability among clusters in Tagalog infixation, and whether these follow the crosslinguistic pattern of section 2.1. The data to be discussed in section 3 come from established loan clusters, and those in section 4 come from poorly attested clusters. In both cases, speakers' treatment of clusters does follow the crosslinguistic pattern.

**3. Corpus.** The first set of data comes from a written corpus of Tagalog. The corpus is made of text from the Web. The method for constructing it was as follows. First, a smaller corpus, generously supplied by Rosie Jones (and derived from Ghani, Jones & Mladenec 2004, which inspired the procedure used here), was used to estimate Tagalog word frequencies. A program generated strings composed of frequent Tagalog words, such as those shown in (8).

(8)	string	glosses
	kami pangulo	'we ( <u>exclusive</u> )' 'president/chief'
	lalo parang	'more/much' 'for- <u>linker</u> '
	tagalog pagiging	'Tagalog' 'being'

noong akin aklat 'then-linker' 'mine' 'book'

A program written by research assistant Ivan Tam sent these strings as queries to Google (www.google.com), using the Google Web APIs service. The service allows a maximum of 1,000 queries per day, with each query returning a maximum of 10 URLs (web page addresses); if a query produces more than 10 results, only 10 are returned at a time and each request for the next 10 counts as another query. Thus, a theoretical maximum of 10,000 URLs can be retrieved per day, but the typical number is approximately 5,000, since not all queries return the full 10 URLs. Because each Google search returns at most 1,000 results, it is important to send a variety of queries in order to give a variety of Tagalog web pages a chance to surface in the top 1,000.

The URLs retrieved each day are compared against those retrieved so far, and the new ones pulled out. Tam's program then retrieves the full text of each of the new URLs, though an existing program such as GNU wget can also be used. The corpus continues to be augmented and refined, but at the time of the numbers reported here it contained 98,607 pages and approximately 20 million words of Tagalog. In a random sample of 100 pages, 24 are blogs, 21 are discussion forums, 13 are newspaper articles, 9 are bible verses, 5 are press releases, 4 are nongovernmental organizations' and social clubs' sites, and the remaining 24 are poetry, articles from sources other than newspapers, book reviews, business and shopping sites, educational materials, glossaries, government sites, political-party sites, song lyrics, and personal ads.

The corpus can be converted into a list of word types, with token frequencies for each. A fragment is shown in (9).<sup>20</sup>

(9) ...

magbabala	33
magbabalak	21
magbabalance	2
magbabalangibog	2
magbabalangkaskas	4
mag-babalangkaskas	1
magbabalanse	2
magbabalaod	10
magbabalat	2
magbabalatkayo	7
magbabalaud	5
magbabalay	2
magba-balebol	1
...	

This file can then be searched for regular expressions corresponding to potentially infix forms, such as `[ptk]in[lr][aeiouwy]` (p, t, or k followed by in, followed by l or r and then a, e, i, o, u, w, or y). The results must be hand-checked to eliminate strings that are not actually infix forms, such as the proper name mckinley.

The initial clusters that have been borrowed into Tagalog unepenthesized are almost exclusively C-glide and stop-liquid.<sup>21</sup> (As discussed in section 4, SC clusters other than s-glide normally undergo prothesis, so that the stem is no longer cluster-initial.) But we can still test one prediction made by Fleischhacker's perceptual account. Although she does not compare different

stop-C clusters, we can compare stop-liquid to stop-glide in the corpus data. Fleischhacker's perceptual explanation predicts that stop-glide should be more splittable than stop-liquid, just as sibilant-glide was found to be more splittable than sibilant-liquid.

Figure 1 and Figure 2 show resulting frequencies for both split and unsplit variants, for both types of cluster (ty, dy are omitted because they can function as digraphs for [tʃ], [dʒ]; reduplicated forms are also omitted—see section 7.1 for examples of reduplicated, infixated forms). Frequencies are a combination of type and token frequency (most of the frequent stems appear with both variants, so type frequencies alone are not informative, and token frequencies alone would cause the results to be dominated by a few frequent types): each stem type contributes a total of 1 unit to the chart, divided between the appropriate unsplit and split columns, according to the token frequency of unsplit and split variants for that type. For example, in Figure 1, for the stem practice, there are 16 tokens total with in, 6 of prinactice/prinaktis (variant spellings) and 10 of pinractice/pinraktis, so the stem contributes 0.4 (6/16) to the CCin (unsplit) column for stop-r, and 0.6 (10/16) to the CinC (split) column for stop-r. Figure 2, for the infix um, works the same way. Percent splitting (using the same token-weighted type frequencies) is also shown for those categories with a sufficient number of types.

INSERT Figure 1 ABOUT HERE

INSERT Figure 2 ABOUT HERE

The main trend to note is that for stop-liquid clusters, nonsplitting is more common, but for stop-glide clusters, splitting is more common. This is true for both infixes, though the

numbers are smaller for um. The trend towards splitting seems to be sharper for stop-glide clusters with um (though overall numbers are smaller). This may be because of a fact observed by Orgun and Sprouse (1999): there is a strong dispreference for the infix um to follow w or m. In the case of Cw clusters, this would mean that there would be an additional pressure for um to split the cluster (and since most of the stop-glide data are from stop-w clusters, this probably explains the difference). Within each of the two charts, the difference between stop-r and stop-w is significant by one-tailed Fisher's Exact Test ( $p < .05$  for in,  $p < .002$  for um).<sup>22</sup> Differences within the same cluster type across the two charts are not significant.

There is a possible etymological confound.<sup>23</sup> English is poor in words beginning with stop-glide sequences (except for Cju, such as few [fju]), and the stop-glide categories in the corpus data are made up entirely of Spanish loans, whereas the stop-liquid categories are a mix of English and Spanish loans. If there is a difference in splitting behavior between the two etymological classes, this could skew the results. Figure 3 and Figure 4 show the results for Spanish-origin<sup>24</sup> loans only, and although the numbers are smaller, the trend remains the same (for stop-w and stop-y, of course, the numbers remain exactly the same). As before, the significant difference within each chart is stop-r vs. stop-w ( $p < .05$  for in,  $p < .005$  for um). There are no significant differences between the two charts. For stop-l and stop-r with each infix, there are no significant differences between Spanish-etymology words and English-etymology words.

INSERT Figure 3 ABOUT HERE

INSERT Figure 4 ABOUT HERE



As a reviewer points out, this lack of Spanish/English difference suggests that fine details of the source-language form are of little importance. Spanish and English  $\underline{r}$  are very different phonetically ([r] vs. [ɾ]), for example, so we might expect CR clusters from the two languages to be treated differently, if speakers are comparing the source-language form to the infix form. (Because Spanish-Tagalog contact is now limited, comparisons to Spanish forms would have had to take place long ago and be fossilized in the word's contemporary behavior.) Spanish and English  $\underline{l}$  are also quite different following a word-initial voiceless stop. We might expect Spanish [pl] to be more splittable, because the [l] in p-in-l... is similar to the original; whereas English [p<sub>0</sub>l] might be less splittable, because the [l] in p-in-l... is somewhat different from the original [l]. The lack of (significant) difference between the two etymological classes is consistent with the proposal that speakers compare the adapted form of the loan—not the source-language original—to the infix form. We might also expect different treatment for Spanish- and English-origin loans if infixation behavior results from variable misperception of cluster-initial foreign words (see section 7.1). The lack of difference suggests that if misperception does take place, it is not much influenced by the phonetic differences between Spanish and English in this case.

There may well be other factors that determine an item's likelihood of splitting (see note 22). It would also be of interest to know whether individual stems have acquired lexicalized behaviors, but there are not enough sets of stems that are identical on all relevant properties to compare. The histogram in Figure 5 shows how many words with a frequency of at least 5 display each rate of splitting. Although the distribution is tail-heavy, suggesting that at least some items behave consistently, it may be that these words simply have properties that especially suppress or promote splitting, and not that they are lexicalized.

INSERT Figure 5 ABOUT HERE

To summarize the corpus results: as predicted, stop-glide clusters are treated as more splittable than stop-liquid clusters. These results are not decisive, however, on the question of whether speakers have implicit phonetic knowledge and a bias in how to apply it. The Spanish loans, especially, have been in the language for some time, so it is possible that rather than individual, on-the-spot decisions about how to infix words, we are now witnessing frozen forms that have been passed down, and that the original motivation for treating stop-glide and stop-liquid clusters differently did not involve any bias on speakers' part. For example, as mentioned in note 19, some older loans from Spanish have an epenthetic vowel, as in palantsa 'iron', from Spanish plancha. If, as appears to be the case (and as would be predicted by Fleischhacker), this epenthesis is more common in stop-glide clusters than in stop-liquid clusters—whether because of judgments of perceptual similarity to Spanish, misperception of Spanish, or some other cause—the greater splittability of the stop-glide clusters, even if part of the synchronic grammar, could be a relic of their previous status as nonclusters and not reflect any phonetic judgments (see section 7.1 for further discussion along these lines).

Tagalog has partial reduplication—marking aspect, among other things—that copies the first C\*V of the stem. Following the crosslinguistic evidence on reduplication discussed by Fleischhacker (recall section 2.1), and extrapolating to the expected splittability difference between stop-liquid and stop-glide clusters, we expect that stop-liquid clusters should simplify less often than stop-glide. Rough corpus counts corroborate this expectation (type frequencies only—not token-weighted, strict matching to  $C_1(C_2)V_1-C_1C_2V_1$  only). For stop-liquid clusters,

there are two options: simplified (e.g. mag-ta-trabaho ‘will work’) and fully copied (mag-tra-trabaho). The simplified option occurs about one third of the time in the corpus and the fully copied two thirds of the time. For stop-glide clusters, there are three options, two simplified and one fully copied. Simplification can occur either by simply skipping the glide (mag-ba-byahe ‘will travel, mag-ke-kwento ‘will narrate’), or, more commonly, by having a reduplicant vowel that corresponds in color to the glide (mag-bi-byahe, mag-ku-kwento). The two simplified options together occur about two thirds of the time, and the fully copied option (mag-bya-byahe, mag-kwe-kwento) one third of the time. Thus, the stop-glide clusters do simplify more often, as expected. But, as with infixation, the fact that these two cluster types have been in the language for some time makes them suspect. Perhaps they reflect epenthesis patterns at an earlier stage (epenthesis more likely into stop-glide than stop-liquid) rather than any phonetic judgment.

A better testing ground, then, would be clusters that are unattested or nearly unattested, since there should be no existing convention on how to treat them, and speakers will be forced to make their own decisions. Such a testing ground does exist: sibilant-consonant (SC) clusters. Except for s-glide, SC clusters are rare word-initially in Tagalog. Spanish does not allow word-initial SC clusters except for s-glide, so no such clusters come in from Spanish loans. English does have a range of SC clusters, but, except for s-glide, they normally undergo prothesis when borrowed into Tagalog. For example, ‘scan’ is normally pronounced [ʔiskan], and the infix is placed before the prothetic vowel ([ʔumiskan], cf. native [ʔumawit] ‘to sing’—see note 28 for discussion of the glottal stop’s status), so that the issue of whether to split the cluster does not arise. Nonprothesized forms such as [skan] are used by some speakers, but they very rarely occur with infixation. In the corpus, there were only 24 tokens, 17 of them from a single type, the nickname of a sports team, which may have originated in a speech error.<sup>25,26</sup>

What will speakers do, then, if forced to perform infixation on words beginning with SC clusters? Will they follow the crosslinguistic pattern identified by Fleischhacker?

**4. Survey.** A survey was conducted to probe speakers' behavior on sibilant-consonant clusters, as well as to confirm the corpus findings on stop-consonant clusters. The survey was conducted over the web. This allowed participants to be located anywhere in the world while completing the survey. It was hoped that many of the participants would be living in the Philippines, and 35 (out of 62 participants who provided usable data) reported that they were. Participants were recruited through announcements in Tagalog-language web forums that contained a link to a welcome page. The welcome page collected demographic information and screened out non-Tagalog speakers (directions and questions were in Tagalog, with responses typed into plain textboxes; understanding of Tagalog was thus necessary to provide appropriate answers). The participant would then see 14 screens like the one shown in Figure 6. Every second item began with a 'fun fact' in teaser-and-answer form (the material at the top of Figure 6, before the forced-choice question). This was the only reward for participation. The materials were real sentences adapted from the corpus. The participant was asked to choose the best option to fill in the blank, and then rate each option. The stimuli were real words when possible, except that any prothetic vowel in the original sentence was removed. For sm and sn, no good examples could be found, so sentences with Tagalog synonyms of smuggle and snow were used, and the loans substituted (without prothesis) for the original words. Item and response orders were randomized separately for each participant. Professional translations were provided by 101 Translations. See the appendix for details on the survey materials and criteria for data inclusion.

INSERT Figure 6 ABOUT HERE

Results are of two kinds, choices and ratings. Figure 7 shows, for each cluster type, the proportion of the time that participants chose the split-cluster option (since this was a binary forced-choice task, the proportion of the time that participants chose the nonsplit option is simply the mirror image). We can see that splitting was seldom chosen for s-stop clusters (on the left), but was usually chosen for sw clusters (on the right).

INSERT Figure 7 ABOUT HERE

It is surprising that participants chose to split sl and shr more often than not, given that stop-l and stop-r clusters were found to split less than 50% of the time in the corpus (see Figure 1, **Error! Reference source not found.**), and Figure 3). Although the survey was not designed to compare s-liquid to stop-liquid, it included stop-liquid filler items. Consistent with the corpus, splitting rates for filler items in the survey are 48% for stop-l (cf. 63% for sl), 43% for stop-r (cf. 77% for shr), and 86% for stop-w (similar to the 90% for sw). Lower splittability for stop-liquid than sibilant-liquid is inconsistent with Fleischhacker's finding that in epenthesis, Farsi prothesizes sl but splits stop-l, and Wolof has variation for sl (and possibly for sr), but only splitting for stop-l and stop-r (see (2)). I have no explanation for this disparity between stop-l/r and sibilant-l/r, except to note that stop-l/r clusters are well attested with infixation among existing loans, whereas infixation of sibilant-l/r is basically novel—perhaps cluster novelty increases (for some unknown reason) the attractiveness of the split option.

Figure 8 shows, for each cluster type, the average rating assigned by participants. Error bars indicate 95% confidence intervals. Note that the vertical axis shows the full range of possible ratings, from 1 (worst) to 7 (best). Looking first at the heavier line with diamonds—CxxC, ratings for split-cluster options—we see that the rating is lowest for s-stop clusters, and highest for sw. The lighter line with squares (CCxx) shows ratings for nonsplit options. Although the rating is highest for s-stop clusters, it is still not very high. This is to be expected, since normally a word beginning with an s-stop cluster would undergo prothesis; that is, neither infixation option is expected to be very acceptable (the survey did not include well-formedness ratings of the stems by themselves).

INSERT Figure 8 ABOUT HERE

In order to determine how much of the ratings pattern is significant, we can compare the rating difference for each pair (split rating minus unsplit rating). Performing a repeated-measures ANOVA with one within-subjects factor (cluster type) with six levels (sT, sm, sn, sl, fr, sw), and no between-subjects factors, with rating difference as the dependent variable,  $F(5.00, 50.00)^{27} = 5.80$ ,  $p < .001$ , with a (vacuous) Huynh-Feldt correction to degrees of freedom. We can then perform paired (by participant)  $t$ -tests on each pair of cluster types. Table 3 shows, for each pair of clusters, whether they behave significantly differently according to each of two measures:  $t$ -test comparison of rating differences between split and unsplit, and Fisher's Exact Test on the number of times the split and unsplit options were chosen in the forced-choice task. Because the crosslinguistic data predict in advance in which direction each difference should be, the  $p$ -values shown for all tests are one-tailed: they test whether there is a difference in the

predicted direction. No differences in the nonpredicted direction (that is, the ratings and choices for sn vs. sl) were significant.

INSERT Table 3 ABOUT HERE

If we take the survey results as supporting a distinction between two clusters if they differ significantly on at least one of the two tasks, we have the four-way distinction shown in Figure 9.

INSERT Figure 9 ABOUT HERE

Although the predicted distinctions among sn, sl, and ʃr were not seen, I conclude from these results that Tagalog speakers do indeed make distinctions among non-sw SC clusters (a three-way distinction, at least), despite having almost no previous experience of how to infix words that begin with them. This suggests that speakers do have implicit knowledge about the splittability of these clusters.

**5. Analysis.** Steriade (2001a, 2001b, 2003) proposes that language users have a P-map, or perceptual map, that they can use to look up the perceptual distance between two fragments of phonological material, such as word-final voiced bilabial stops vs. word-final bilabial nasals. Steriade argues that these P-map distances translate into constraint rankings: a faithfulness constraint is ranked by default according to the size of the perceptual difference that its violation creates. That is, if constraint FAITH1 is violated when underlying x becomes surface y, and FAITH2 is violated when underlying z becomes w, and  $\Delta(\underline{x}, \underline{y}) > \Delta(\underline{z}, \underline{w})$ , then, by default FAITH1

>> FAITH2 (for underlying-surface or input-output correspondence—the same principle applies within other correspondence-constraint families, such as output-output or base-reduplicant.) If a learner has no language-specific evidence to overturn that ranking, then the ranking stands, though it may be detectable only through probes such as literary invention, loan adaptation, and experimental tasks; it is possible, however, that a series of diachronic changes could lead to a situation in which the data compel learners to overturn the default ranking.

The similarity hierarchy proposed by Fleischhacker 2000a (4) is repeated as (10), with  $\underline{S}$  substituted for  $\underline{C}_1$  (and all distinctions treated as real). Adopting Steriade's proposal, Fleischhacker translates the similarity scale into the constraint ranking in (11).

$$(10) \Delta(ST, SV) > \Delta(Sm, SV) > \Delta(Sn, SV) > \Delta(Sl, SV) > \Delta(SR, SV) > \Delta(SW, SV)$$

$$(11) \text{DEP-V/S}_T \gg \text{DEP-V/S}_m \gg \text{DEP-V/S}_n \gg \text{DEP-V/S}_l \gg \text{DEP-V/S}_R \gg \text{DEP-V/S}_W$$

DEP constraints (McCarthy & Prince 1995) penalize insertion of segments. These are context-sensitive DEP-V constraints, which penalize insertion of a vowel in a particular context, such as between a sibilant and a stop ( $S\_T$ ) as in /sparta/ → [separta]. By ranking LEFT-ANCHOR (McCarthy & Prince 1995: the leftmost segment of the underlying form must correspond to the leftmost segment of the surface form) at some point in this scale, Fleischhacker obtains a given language's cut-off point for cluster splitting. Additional markedness and faithfulness constraints determine which unsplit clusters are adapted faithfully and which receive a preceding epenthetic vowel. Prince & Smolensky's 1993/2004 \*COMPLEX, a markedness



constraint that penalizes, among other structures, initial consonant clusters, drives the epenthesis. For languages where no clusters receive a preceding epenthetic vowel, the cut-off constraint is not LEFT-ANCHOR but rather a markedness constraint against consonant clusters. The tableaux in (12) illustrate the analysis for a language which prothesizes sibilant-stop clusters, and epenthesizes sibilant- $\_l$  clusters.

(12) Schematic analysis of asymmetric epenthesis pattern

	source word [spV...]	*COMPLEX	DEP-V/S__T	LEFT-ANCHOR	DEP-V/S__l
a.	spV...	*!			
b.	sipV ...		*!		
c.	☞ ispV ...			*	

	source word [slV...]	*COMPLEX	DEP-V/S__T	LEFT-ANCHOR	DEP-V/S__l
d.	slV ...	*!			
e.	☞ silV ...				*
f.	islV ...			*!	

In order to extend this account to similar patterns in reduplication, imperfect puns, and alliteration, Fleischhacker (2000b) introduces an additional family of default-ranked contextual MAX constraints, which penalize deletion of segments (McCarthy & Prince 1995), shown in (13) (cf. Fleischhacker's DEP family in (11)). In reduplication, the relevant constraint for splitting is not DEP but MAX, since a segment of the base is deleted in the reduplicant (gaí-grōt). In

imperfect puns and alliteration, the relevant constraint is either DEP or MAX, depending on which member of the pair is taken as primary (**Bonaparte/Blown-apart**).

(13) MAX-T/S\_V >> MAX-m/S\_V >> MAX-n/S\_V >> MAX-l/S\_V >> MAX-R/S\_V >>  
MAX-W/S\_V

To further extend the account to infixation, neither DEP nor MAX suffices, since there is no epenthesis or deletion involved. The faithfulness constraint that is violated by infixation within a cluster is CONTIGUITY (McCarthy & Prince 1995), which requires adjacent segments' correspondents to remain adjacent. In the context-sensitive CONTIGUITY family in (14), particular consonant clusters in the uninfixed form are required to remain adjacent in the infixed form.

(14) CONTIG-ST >> CONTIG-Sm >> CONTIG-Sn >> CONTIG-Sl >> CONTIG-SR >> CONTIG-SW

This is not quite right, however, because the ranking in (14) follows from the similarity hierarchy in (10) only if the reason for the contiguity violation is insertion of material beginning with a vowel, as in infixation or vowel epenthesis. For example, (10) says nothing about which pair is more similar, (st, spt) or (sl, spl), but (14) says that inserting p into st (violating CONTIG-ST) is worse than inserting p into sl (violating CONTIG-Sl). If we want the contextually sensitive CONTIG family to reflect the similarity claims in (10), we must further specify the context in which the CONTIGUITY constraint applies, as in CONTIG-ST/V..., meaning 'adjacent ST in one

form must not have their correspondents in another form separated by a string beginning with a vowel':

(15) CONTIG-ST/V... >> CONTIG-Sm/V... >> CONTIG-Sn/V... >> CONTIG-Sl/V... >>  
 CONTIG-SR/V... >> CONTIG-SW/V...

To avoid excessive digression, I adopt this approach, but point out the possibility that (11), (13), and (15) could be unified under a more general type of constraint, \*MAP-S<sub>1</sub>S<sub>2</sub>(<sup>A</sup>X<sup>B</sup>,<sup>C</sup>Y<sup>D</sup>): an X in the environment A\_\_B in string S<sub>1</sub> must not correspond to a Y in the environment C\_\_D in string S<sub>2</sub>. For example, \*MAP(S<sup>T</sup>,S<sup>V</sup>) forbids a sibilant that is followed by a stop from corresponding to a sibilant that is followed by a vowel. The hierarchy \*MAP(S<sup>T</sup>,S<sup>V</sup>) >> \*MAP(S<sup>m</sup>,S<sup>V</sup>) >> \*MAP(S<sup>n</sup>,S<sup>V</sup>) >> \*MAP(S<sup>l</sup>,S<sup>V</sup>) >> \*MAP(S<sup>r</sup>,S<sup>V</sup>) >> \*MAP(S<sup>w</sup>,S<sup>V</sup>) would cover the three hierarchies in (11), (13), and (15). Donca Steriade (p.c.) suggests a less radical move that would also cover all three cases: C-CONTIGUITY(S,T), defined as 'if S precedes T in one form, the correspondent of S in the other form must not be followed by a vowel (and likewise for other consonant pairs)'. The crucial point here is that faithfulness constraints must be made context-sensitive; less crucial is the point that we can formulate a single family of constraints that covers all the cases (epenthesis, infixation, partial reduplication, etc.).

The CONTIGUITY analysis of infixation is illustrated in (16), which can be compared to (12). The tableaux show an idealized situation in which sibilant-stop clusters never split and sibilant-l clusters always split. Instead of \*COMPLEX, the constraint driving splitting here is ANCHOR-STEM, which requires a word to begin with stem material and thus forces the infix inwards. LEFTMOST (Prince & Smolensky 1993/2004), which keeps the infix as close to the left

as possible, favors splitting. The reason for using ANCHOR-STEM to force infixation rather than Prince and Smolensky's NOCODA is that infixation within a cluster is not predicted under their analysis, since the candidate g-um.-rad.wet has just as many codas as prefixed

\*um.-grad.wet.<sup>28,29</sup>

(16) Schematic analysis of asymmetric infixation pattern

	<u>in</u> + <u>spin</u>	ANCHOR-STEM	CONTIG-ST/V...	LEFTMOST	CONTIG-SI/V...
a.	<b>in</b> spin	*!			
b.	<b>sin</b> pin		*!	s	
c.	☞ <b>spin</b> in			sp	

	<u>um</u> + <u>slip</u>	ANCHOR-STEM	CONTIG-ST/V...	LEFTMOST	CONTIG-SI/V...
d.	<b>um</b> slip	*!			
e.	☞ <b>sum</b> lip			s	*
f.	<b>slum</b> ip			sl!	

We have seen in the corpus data that there actually is variation for every cluster, and the same is true in the survey data. Variable constraint ranking, along the lines of Boersma 1997 and 1998, Hayes & MacEachern 1998, and Boersma & Hayes 2001 can model this. The ranking values shown in Table 4, crafted to obtain the desired pattern (using Hayes & al. 2003),<sup>30</sup> derive idealized outputs shown in Figure 10 (cf. the survey results in Figure 7).

INSERT Table 4 ABOUT HERE

INSERT Figure 10 ABOUT HERE

**6. Discussion of alternatives.** It has been argued above that the survey results on SC clusters can be accounted for by assuming that speakers have implicit knowledge of how the similarity between  $C_1C_2$  and  $C_1V$  varies depending on  $C_2$ , and that they apply this knowledge so as to maximize the similarity of infixed and uninfixed words.

As mentioned in section 1, the structure of the argument here parallels Pullum and Scholz's (2002) definition of argument from poverty of the stimulus. Pullum and Scholz lay out a form of the argument that contrasts language learning using 'inborn domain-specific linguistic information' with learning using 'generalization from experience by the ordinary methods that are also used for learning other (nonlinguistic) things from experience' (p. 17).<sup>31</sup> For our purposes, we can partition the set of possible learning theories differently: on one side are those that endow the learner with a specific phonetic predisposition—in this case, a bias for preserving perceptual similarity between related forms and a way of assessing similarity (the bias and assessment mechanism being possibly language-specific or possibly instances of more general mechanisms)—and on the other side are all other theories, including those that give the learner no language-specific prior knowledge or disposition and those that give the learner some language-specific prior knowledge or disposition, but no preference for maintaining perceptual similarity between related forms. As Pullum and Scholz set it out, then, the researcher's task is as in (17) (cf. Pullum & Scholz's (4)):

- (17) a. Describe what speakers are alleged to know: a constraint ranking like that in (15), or some other means of deriving the observed differences in sC splittability by infixes.
- b. Describe the hypothetical data that, if available, would have allowed learners to discover (a) without any prior knowledge or bias: enough infixes of the form s-in-C.../sC-in-... or s-um-C.../sC-um-... to establish the rate at which the infix splits each type of sC cluster, with sw splitting the most, then shr, sl, sn, then sm, then sT.
- c. Give ‘reason to think’ (p. 19) that, without the prior knowledge in question, learners could not have discovered (a) without (b): sections 7 and 8
- d. Give evidence that learners do not in fact have access to (b): the extremely low corpus frequency of infixes sC-initial stems (except sw), since these forms normally undergo prothesis
- e. Give evidence that learners do nevertheless acquire (a): the survey data.

Pullum and Scholz suggest that (17c) might be accomplished mathematically, using formal learning theory. It is not clear what learnability framework would be appropriate here, where the target language includes as grammatical all the relevant forms, but produces them with different frequencies. Instead, the following two sections construct and evaluate a variety of accounts that do not rely on endowing the learner with a similarity bias. I consider accounts that endow learners with as little language-specific knowledge as possible, though some are not entirely domain-general, relying on distinctive features, for example (section 7), and accounts that endow learners with language-specific knowledge other than a similarity bias (section 8).

Although the proposal here about (17a) differs from Broselow's (1992b), the reasoning is similar. She argues that even speakers of languages that lack initial ST and TL know that these onsets must have different structures, and therefore apply different epenthesis strategies to the two cluster types when learning languages that have them.

**7. Explanations without implicit knowledge?** Is it possible to account for the survey results without attributing implicit phonetic knowledge to speakers? An account based on pure misperception of an infix's location seems implausible—speakers would have to actually mishear kw-in-ento as k-in-wento ('to narrate'), and moreover do so more often than they mishear dr-in-owing as d-in-rowing ('to draw') (or vice versa: mishear k-in-wento as kw-in-ento less often than d-in-rowing as dr-in-owing). But even if such mishearing were possible, it would not account for the survey data, since the SC clusters are ones that speakers have almost never heard with an infix before—there has been (almost) nothing to mishear, and the survey participant must make a decision on the spot.<sup>32</sup>

**7.1. Excrescent vowels.** A more plausible misperception-based account is suggested by the possibility of excrescent vowels, though I present some evidence below that argues against it. Suppose that clusters are splittable to the extent that they are actually pronounced or perceived with an extra vowel. That is, if slip 'slip' is really disyllabic [silip], it should of course be infixed [s-um-ilip].<sup>33</sup> Speakers might still spell the words as slip and sumlip, but they would be treating the stem as though it begins with CV, not with a C<sub>1</sub>C<sub>2</sub> cluster.<sup>34</sup> To explain cluster differences, we could plausibly assume that greater sonority of C<sub>2</sub> encourages the production or perception of an extra vowel, though assumption may itself require an appeal to substantive bias. (See

discussion of Hall's svarabhakti hierarchy in section 8.1). Assuming that these 'extra' vowels have the same status as other vowels, this theory predicts that words with split clusters are treated as though they had an unspelled extra syllable. That prediction is contradicted by some data on infixation with reduplication.

In native words, when infixation and one-syllable reduplication combine, indicating incomplete realis aspect, the result is a prefixed copy of the stem's CV, with an infix after the copied C, as in b-um-a-bago 'is changing', from the stem bago 'new'. When this construction is applied to a cluster-initial loan, several variants are possible. Examples are shown in Table 5, with corpus frequencies.

INSERT Table 5 ABOUT HERE

Variant II, with the onset copied and split, demonstrates that a cluster can be split without being treated as though it has an extra, unspelled syllable (though this variant is, admittedly, not very frequent). If there were such an extra syllable, the variant-II spellings would indicate the pronunciations [g-um-**u**wa-g**u**wapo], [p-in-**o**ro-p**o**roblema], etc., with the first two syllables of the stem copied, which is inconsistent with the general reduplication pattern. Accommodating variant II under the excrescent-vowel account requires the putative excrescent vowel to have an intermediate status: like Hall's svarabhakti cases, the excrescent vowel is ignored prosodically in syllabification; but unlike Hall's cases, the excrescent vowel would still count as a segment, in order to condition infixation.



**7.2. Cluster frequencies.** Another possible explanation for the survey results is based on initial cluster frequencies.<sup>35</sup> Consider the possibility that speakers, using knowledge of English, are able to identify instances of prothesis, as in iskor ‘score’, and that they interpret prothesis as evidence of a cluster’s nonsplittability, since anaptyxis (\*sikor) can be observed not to have occurred. Then, the word-initial SC clusters of English loans that are observed with a prothetic vowel most often might be treated as the least splittable. Under this account, speakers would have implicit knowledge of splittability, but that knowledge would not be phonetic and would be based on direct evidence. Corpus data can be used to evaluate the viability of this possibility. In order to keep the amount of data to be inspected manageable and to minimize the number of spurious items, counts are restricted to prothesized English loans beginning with SC clusters that carry some Tagalog morphology (reduplication, infixation, prefixation, and/or suffixation). The counts in Figure 11 do show that ST clusters appear most often, which could explain their low level of splittability. But the greater splittability of sn compared to sT is not explained, since sn is about as frequent as sp, st, and sk. The prediction for a sm-sn difference is in the wrong direction: since sn is much more frequent than sm, it should be less splittable, not more splittable as it was in the survey. The frequency idea has nothing to say about differences between TR and TW, since neither undergoes prothesis.<sup>36</sup>

INSERT Figure 11 ABOUT HERE

**7.3. ‘Insert infix before first X’** Finally, a reviewer suggests an explanation that extrapolates from the distribution of um and in in native words. Two facts about native words must be introduced. First, the infix in has an allomorph, the prefix ni. When no other prefix is present, the

prefix ni occurs variably for stems that begin with l, w, y ([j]), and h (e.g. ni-luto~ l-in-uto ‘cooked’), predominating over infixation for l and y, and rare for h and w. When the prefix [ʔi] is also present, the frequency of ni increases for stems beginning with those consonants, and ni is also used, obligatorily, with stems that begin in glottal stop (or vowel, if the glottal stop is viewed as epenthetic): [ʔi-ni-ʔabot] ‘handed to’, cf. [ʔ-in-abot] ‘reached for’; in is still used, however, for other consonants (e.g. [ʔi-b-in-uhos] ‘poured into’). Second, the infix um rarely occurs at all with stems beginning in m, w; most such stems simply lack an um form (see Orgun & Sprouse 1999).

To project these facts onto new words, assume that Tagalog speakers formulate and evaluate the reliability of generalizations of the form ‘insert the infix before the first X’, where X is a feature matrix.<sup>37</sup> The simplest and most reliable generalization has X=[+syllabic] (‘insert the infix before the first vowel’), but others have some support too. X=[-consonantal], for example, does fairly well, because it is true of words like b-um-ili ([i] is [-cons]), and also of words like ni-yakap (y=[j] is [-cons]), and false only of words like y-um-akap (the rarity of words like w-um-agayway helps boost this generalization). X=[+sonorant] does less well, but still not too badly. It is true of words like b-um-ili and ni-lagay, but false of words like l-um-akad. One of the worst generalizations, with X=[-sonorant], is never true (assuming [h] and [ʔ] are [+sonorant]): ni never occurs with obstruent-initial stems. If speakers then apply these generalizations in the survey task, we can see that s-in-werte, which obeys ‘insert the infix before the first [-cons]’ should be rated higher than s-in-lip, which obeys the weaker ‘insert the infix before the first [+son]’, which should in turn be rated higher than s-in-top, which obeys the always-false generalization ‘insert the infix before the first [-son].’

There are many values for  $\underline{X}$ , however—as many as there are natural classes in the phoneme inventory—and in order to make the account truly data-driven, the learner must consider all of them. In the Tagalog case, using a fairly standard feature set, there are 786 natural classes. A useful framework for evaluating a large set of constraints/generalizations set is the Maximum Entropy framework. A full explication of this framework is beyond the scope of this paper; see Goldwater & Johnson 2003 on applying Maximum Entropy to OT-like constraint weighting. In the course of learning, each generalization is assigned a weight (or in this case, the inverse of each generalization is assigned a negative weight). When it comes time to generate a form, the log probability that any given candidate is chosen is the weighted sum of its constraint violations. Learning proceeds by adjusting weights to maximize the likelihood of the training data. Here, the Conjugate Gradient Method (Hestenes & Stiefel 1952; see Shewchuk 1994 for a tutorial) was used.<sup>38</sup> The training data—the frequency of  $\underline{ni}$ ,  $\underline{in}$ , and  $\underline{um}$  for stems beginning in  $\underline{CV}...$ , for all values of  $\underline{C}$  and  $\underline{V}$ —were derived from a mixture of corpus counts and extrapolations from a database of disyllabic native roots, from English 1986.<sup>39</sup>

As shown in Figure 12, the resulting set of weighted generalizations manages to distinguish  $\underline{sw}$  from the other  $\underline{sC}$  clusters, splitting it about 5% to 50% of the time, depending on the following vowel (because different vowels belong to different sets of natural classes), but incorrectly predicts that the other clusters should all split less than 10% of the time:

INSERT Figure 12 ABOUT HERE

At present, I conclude that none of the data-driven accounts of the survey data works well enough to be accepted. Of course, it remains to be seen what others can be devised.

## 8. Other candidates for implicit knowledge

**8.1. Excrescent vowels II.** An alternative to the perceptual account given in section 5 might be an articulatory account. Hall (2003) proposes that svarabhakti vowels (vowels, sandwiched between two consonants, that do not contribute to the syllable count, and that have either the same quality as a nearby vowel or a default quality), which she proposes are articulatorily distinct from true epenthetic vowels, result from loosely coordinated consonant articulations. If two adjacent consonants are pronounced with a gap in between—that is, the first consonant’s closure is released before the next consonant’s closure begins, so that there is a short interval in which the vocal tract is open—an excrescent (svarabhakti) vowel is perceived, although no actual segment has been inserted. If an adjacent vowel’s gesture overlaps that gap, the excrescent vowel has the same quality as that adjacent vowel; otherwise, the excrescent vowel has a default quality, such as schwa. An example from Hall is Dutch [kələm], a variant of [kəlm] ‘calm’.<sup>40</sup>

Hall examines the distribution of svarabhakti vowels crosslinguistically and finds many regularities. First, these vowels occur only when at least one of the surrounding consonants is a sonorant. Hall attributes this to the relative unmarkedness of vowel-sonorant overlap (compared to vowel-obstruent overlap) and to special phasing constraints for sonorants that cause them to be more loosely coordinated with other consonants. In both cases, the reason for sonorants’ special behavior is unknown. It might be articulatory, but, as Hall discusses, it might be perceptual: there is a body of phonetics research arguing that gestures are timed so as to maximize their perceptual recoverability (Kingston 1990, Silverman 1997, Wright 1996, and many others), and it may be that sonorants, especially in the V\_\_C environment that Hall focuses on, are more perceptually vulnerable than other consonants.



Tagalog survey data are consistent with Fleischhacker's hierarchy and not with a splittability interpretation of Hall's, suggesting that loosely coordinated articulation is not the source of splittability. Still, Hall's evidence for putting glides to the left of liquids in this hierarchy comes only from Hausa; most of the languages she surveys lack glides in the relevant environment. The other differences are less definite. First, there are no loanwords beginning with a C-guttural cluster in Fleischhacker's survey (and a source language providing such words would be hard to find), so gutturals do not appear in her hierarchy. And second, Fleischhacker groups all rhotics together. The two languages in her survey that distinguished laterals from rhotics were Farsi and (Bharati's) Hindi. In Farsi, where S-rhotic clusters are split but Sl clusters are not, the rhotic is a tap, [r] (seri laŋka 'Sri Lanka', Shabnam Shademan, p.c.), which would not be a mismatch with Hall's hierarchy. In Hindi, where, in Bharati's description, S-rhotic clusters are split but Sl vary, the rhotic is presumably a trill, which would be a mismatch with Hall's hierarchy. The rhotic in the Tagalog cases can be a tap, which both hierarchies (and the survey data) put to the right of laterals, or an approximant [ɹ], which does not occur in the languages examined by either Hall or Fleischhacker (or, rarely, the Tagalog rhotic can be a trill).

One can imagine an extension of an articulatory-splittability account to reduplication. For alliteration and puns, we would need to assume that the appropriateness of an alliteration or pun is judged in articulatory terms. Whether an articulatory account has anything to say about Fleischhacker's similarity-judgment experiment is the least clear; we would have to suppose that subjects listening to a stimulus pair are not merely comparing them perceptually, but are perhaps comparing them as articulatory variants.

Overall, it is unclear whether Hall and Fleischhacker offer potentially competing accounts of the same range of phenomena—with some discrepancies to be resolved—or accounts

of different phenomena that happen to result in largely overlapping cluster-splittability scales (with Tagalog following Fleischhacker's scale). If the former, I lack evidence to determine whether the implicit knowledge demonstrated by Tagalog speakers in the survey task is perceptual or articulatory.

**8.2. Destruction of marked clusters.** Another alternative to the perceptual account is that speakers' implicit knowledge does not concern cluster splittability at all, but concerns the markedness of the infixes. One possibility is that speakers deploy infixes so as to eliminate marked clusters. We would therefore expect that marked clusters would split the most often, and unmarked clusters would split the least often. This seems, however, to be the opposite of what happens. The splittability hierarchy is repeated in (20) with grouping into broad sonority classes, and it seems that the clusters that split the least often are actually the most marked, and vice versa.<sup>43</sup>

(20) <u>least often split</u>	<u>most often split</u>
sibilant-stop ( <u>ST</u> )	sib.-nasal ( <u>Sm</u> , <u>Sn</u> )
sib.-liquid ( <u>Sl</u> , <u>Sr</u> )	sib.-glide ( <u>SW</u> )
stop-liquid ( <u>Tl</u> , <u>Tr</u> )	stop-glide ( <u>TW</u> )
<u>most marked</u>	<u>least marked</u>

There are a few criteria we could use to determine which clusters are more marked. Crosslinguistically, it has been claimed that the greater the sonority increase from  $C_1$  to  $C_2$ , the less marked is the onset cluster  $C_1C_2$  (e.g. Greenberg 1978, Selkirk 1984).<sup>44</sup> This would mean that TW is less marked than T-liquid and that the SC clusters towards the right in (20) are less

marked than those towards the left. Steriade's (1999) theory of consonant cuing claims that consonant clusters are marked because of  $\underline{C}_1$ 's reduced perceptibility:  $\underline{C}_1$  lacks a following vowel or sonorant whose formants it can alter, and lacks a release burst. This predicts that greater sonority of  $\underline{C}_2$  should reduce markedness: again, TW should be less marked than T-liquid, and that the SC clusters towards the right should be less marked than those towards the left. Under both theories of markedness, it is actually the more marked clusters that split the least often.

Tagalog-internal evidence, though limited, points in the same direction: more marked clusters split less often. We can look first at adaptation of English loans, where T-liquid, TW, and SW are freely tolerated, but not other word-initial SC clusters. (They are, as discussed in section 4, typically repaired by prothesis.) This would suggest that T-liquid, TW, and SW are less marked than the rest, even though they split the most often. Second, within native words, there is often variation between  $C_1VC_2$  and  $C_1C_2$  when  $C_2$  is a glide (and  $V$  matches it in color, i.e. backness and rounding), but not when  $C_2$  is a liquid, no matter what the intervening vowel:<sup>45</sup>

- (21) [piják] ~ [pják] 'squawk'  
 [buwán] ~ [bwán] 'moon'  
 [purók] \*[prók] 'district'

This suggests that TW is less marked than T-liquid, though it is also possible that similarity preservation is at work here. That is, since TVW is highly similar to TW—especially if  $\underline{V}$  matches the glide in color—deletion of  $\underline{V}$  is permissible, but since T-V-liquid is less similar to T-liquid, deletion is not permissible there.



A final piece of Tagalog-internal evidence that TW are the least-marked clusters (even though they split the most often) comes from syllabification. Word-internal clusters are normally syllabified C<sub>1</sub>.C<sub>2</sub>, avoiding a complex onset. Evidence for this syllabification comes from speakers' intuitions (Schachter & Otones 1972) and from stress facts. Stress (sometimes characterized as length—see Schachter & Otones, French 1988, and Zhang 2001 for discussion) in native Tagalog words can fall on either the penult or the ultima, except not on a closed penult (22a). When a verbal suffix is attached, stress shifts one syllable to the right (22b).

(22) (a) unsuffixed forms	(b) suffixed forms
Open penult: penultimate or final stress	
[bí.roʔ] 'joke'	[bi.rú.ʔ-in] 'to joke'
[ta.nónŋ] 'question'	[ta.nu.ŋ-ín] 'to question'
Closed penult: final stress only	
[ʔik.líʔ] 'shortness'	[ʔik.li.ʔ-án] 'to shorten'

(English 1986)

Loans can have stress on a closed penult (23), but these words behave differently under suffixation: stress shifts to the final syllable (with secondary stress sometimes remaining on the closed syllable), as shown in (23a). There are some rare exceptions to this pattern that behave as though the penult were not closed—stress shifts one to the right (23b). Those cases all involve a C-glide cluster. Apparently, word-internal C-glide clusters can optionally be syllabified as complex onsets, suggesting that C-glide is less marked as an onset than other types of cluster. Again, this makes the wrong prediction for the splitting facts.

- (23) a. [rén.da] ‘rein’ [ren.da.h-án] ‘to rein’ (Spanish rienda)  
 b. [di.lí.rjo] ‘delirium’ [di.lì.-di.li.rjú.-han] ‘feigned delirium’ (Spanish delirio)

(English 1986)

**8.3. Avoidance of marked clusters.** A second markedness-based possibility is that speakers are avoiding the creation of marked clusters. Whenever a CC cluster is split by a VC infix, a new cluster is created, as the mr cluster of g-um-raduate. If this force is responsible for differences in cluster splittability, then we expect that  $C_1C_2$  should be more splittable the less marked a nasal- $C_2$  cluster is. Again, this is the opposite of what happens:

- |                                 |             |              |                           |
|---------------------------------|-------------|--------------|---------------------------|
| (24) <u>least often created</u> |             |              | <u>most often created</u> |
| nasal-stop                      | nasal-nasal | nasal-liquid | nasal-glide               |
| <u>least marked</u>             |             |              | <u>most marked</u>        |

In order to establish nasal-C cluster markedness, we can look at both crosslinguistic and Tagalog-internal evidence. Vennemann’s (1988) crosslinguistically based Syllable Contact Law posits that coda-onset transitions should be of falling sonority. That would make nasal-stop the least marked cluster. If we interpret the syllable contact law gradiently, so that flat sonority is also worse than rising sonority, and that the greater the sonority rise, the worse, then the clusters in (24) become more marked towards the right.

Tagalog-internally, we can compare type frequencies of root-internal nasal-C clusters, shown in Figure 13.<sup>46</sup> Nasal-stop clusters have the highest raw frequency (see bar labeled ‘nt,

etc.’), as well as the highest frequency relative to the control case, oral-stop clusters (‘kt, etc.’). By those criteria, nasal-stop clusters should be the least marked, despite being created least often by infixation. (All three Tagalog nasals are combined since their postnasal frequency is so low; there is no column for  $C_2=r$ , because [r] in native words does not occur after a nasal.)

INSERT Figure 13 ABOUT HERE

In analyzing loanword epenthesis, Fleischhacker (2005) does appeal to markedness considerations in order to account for the full typology of languages that tolerate some initial clusters and repair others. There does not seem to be a role, however, for cluster markedness constraints in determining infix location in Tagalog (though there may be a role for other markedness constraints, such as Orgun & Sprouse’s (1999) \*w-um- constraint).

**9. Summary.** According to the framework laid out for linguistic investigation in Chomsky 1964, an explanatorily adequate linguistic theory should correctly predict which grammar a learner arrives at after exposure to data. Determining which grammar the learner does arrive at (the descriptively adequate grammar), out of all the grammars that could account for the learning data (the observationally adequate grammars) is itself difficult: we cannot distinguish the descriptively adequate grammar from other observationally adequate grammars merely by inspecting the same data that the learner has access to, because by definition the learning data are consistent with all observationally adequate grammars. This is a particular problem in phonology, because the data we work with are, in many cases, most likely part of the learning data—pronunciations of individual words, for example.

We can use new tasks to establish how speakers generalize beyond a list of memorized items, and thus get a better picture of the descriptively adequate grammar, but this does not directly help us understand cross-linguistic trends. For example, Berko (1958) showed that English speakers can generalize from existing plural nouns to form new plurals by adding [ɪz] after sibilants, [s] after voiceless non-sibilants, and [z] otherwise. This tells us that an explanatorily adequate theory should predict that learners extract this generalization rather than, say, the generalization that [s] is added to form plurals (with many listed exceptions: dog[z], dish[ɪz], feet, etc.). But Berko's results do not tell us whether learners privilege an assimilatory pattern like English's over a dissimilatory pattern—say, [z] after voiceless sounds and [s] after voiced—because a dissimilatory grammar is simply not on the table for learners of English. Cross-linguistic trends are relevant to developing an explanatorily adequate theory only if they tell us something about learner preferences, and as discussed in section 1, Blevins (2004) and others have cast doubt on the assumption that they do.

This study therefore adds to the body of research cited in section 1 that investigates the expectations that humans bring to phonological learning by putting speakers in situations where they are not constrained by the learning data. I have argued that Tagalog speakers are free to learn grammars in which, for example, st onsets are more splittable by infixation than sl, or equal in splittability—yet, as a group, the study participants agree that st is less splittable than sl. More generally, the corpus and survey data presented here have shown that Tagalog speakers' treatment of word-initial clusters parallels the crosslinguistic treatment of these clusters found by Fleischhacker (2000a, 2000b): the more sonorous the second member of the cluster, the more likely that the cluster will be split in such a way that the first consonant becomes prevocalic. The survey data show Tagalog speakers making distinctions even among word-initial clusters that are almost unattested with infixation, making it unlikely that speakers' decisions are based on prior experience of an established convention. I have argued that Tagalog speakers must have some implicit knowledge about these clusters, plausibly how similar the C<sub>1</sub>-C<sub>2</sub> transition is to a C<sub>1</sub>-V transition. Additionally, speakers must have a bias about how to apply that knowledge: the beginning of the infixed form should be similar to the beginning of the uninfixed form.

I also hope to have shown, in section 7, that determining whether speakers are in fact constrained by prior experience is not straightforward. Direct evidence on how to infix SC clusters is scarce, but, depending on our theory of the learner, there are various sources of indirect evidence that could have shaped the grammar. I was unable to construct an account in which indirect evidence could explain the survey data, but this does not rule out the possibility that such an account exists.

## **Appendix: survey details**

### **Instructions**

The welcome page says (in Tagalog)

Thank you for visiting.

This website is a project by Kie Zuraw, assistant professor in the Department of Linguistics at UCLA. Its purpose is to investigate how Tagalog speakers form certain words.

Please participate in the study only if, in your opinion, Tagalog is your native language.

You will be shown a series of 14 sentences taken from informal Tagalog writing, each with one word left out. You will be asked to click on what you think is the best way to fill in the blank. Some of the sentences may use slang or informal grammar and spelling. Please try not to worry about whether the sentence as a whole is correct or not—just decide which is the best way to fill in the blank.

You will also be asked to give a score to each choice by clicking on a number from 1 (worst) to 7 (best).

Depending on the speed of your internet connection, the study should take about 10 minutes to complete.

After every second example, you will see an interesting fact about Tagalog. The first: Do you know what these words have in common: akala, asal, asam, kubol, hukom, halal, hamak, hikayat? Complete the next two examples to see the answer.

This is followed by collection of anonymous demographic data, an option to enter an e-mail address and be notified of future studies, and standard information about the rights of human subjects.

Each item contains, on one page, a forced-choice task and a rating task (see Figure 6 for a sample). For the first item, the forced-choice instructions (in Tagalog) are:

Choose the best word to fill in the blank by clicking the circle next to it. There are no right or wrong answers. We just want to know what, in your opinion, is the best choice.

and those for the rating task are:

Now rate each choice on a scale from 1 (worst) to 7 (best) by clicking the rating you want.

For subsequent items, the directions are abbreviated to

Choose the best word to fill in the blank:

Rate each item from 1 to 7

The rating scale, however, continues to label 1 as ‘worst’ and 7 as ‘best’.

## Materials

Each participant sees fourteen items, with the order randomized for each participant. Six items are SC clusters, and the rest can be considered fillers from the perspective of this study.

target items

- 1 of {in+scan, um+skor, in+specify, in+stop}
- in+smuggle
- um+snow
- um+slip
- um+shrink
- 1 of {in+swerte, um+sweldo}

filler items

- in+byahe
- um+byahe
- in+bwisit
- um+bwelo
- 1 of {in+flash, in+frame}



- 3 of {in+syuting, in+pwesto, in+block, in+break, um+drive, in+drive, in+drowing, um+grabe, um+gwapo, in+create, in+kwento, in+plano, in+promote, in+pwersa, um+pwersa, in+trabaho, um+trabaho}

The two response options are in random order on each trial.

### **Criteria for data inclusion**

A data triple (binary choice plus rating of each option) was excluded if the option chosen received a lower rating than the option not chosen. If a participant made more than 2 such errors, or if the participant completed fewer than 5 items, all data from that participant were excluded.

## References

- Andersen, Henning 1972. Diphthongization. Language 48.11-50.
- Anderson, Stephen 1981. Why phonology isn't 'natural'. Linguistic Inquiry 12.493-539.
- Avery, Peter, and Greg Lamontagne 1996. A note on Tagalog infixation. Abstract from West Coast Conference on Formal Linguistics XV, University of California, Irvine.
- Bach, Emmon, and Robert Harms 1972. How do languages get crazy rules? Linguistic Change and Generative Theory, ed. by Robert Stockwell and Ronald Macauley, 1-21. Bloomington, IN: Indiana University Press.
- Berko, Jean 1958. The child's learning of English morphology. Word 14.150-177.
- Bharati, Surabhi 1994. Aspects of the phonology of Hindi and English. New Delhi: Arnold.
- Blevins, Juliette 2004. Evolutionary phonology: the emergence of sound patterns. Cambridge: Cambridge University Press.
- Blevins, Juliette, and Andrew Garrett 1998. The origins of consonant-vowel metathesis. Language 74.508-556.
- Blevins, Juliette, and Andrew Garrett 2004. The evolution of metathesis. In Hayes, Kirchner & Steriade, 117-156.
- Boersma, Paul 1997. How we learn variation, optionality, and probability. Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam 21.43-58.
- Boersma, Paul. 1998. Functional Phonology. Formalizing the interactions between articulatory and perceptual drives. University of Amsterdam dissertation. The Hague: Holland Academic Graphics.
- Boersma, Paul, and Bruce Hayes 2001. Empirical tests of the Gradual Learning Algorithm. Linguistic Inquiry 32.45-86.

- Breen, Jim n.d. Jim Breen's WWWJDIC Japanese-English Dictionary Server. Web site, <http://www.csse.monash.edu.au/~jwb/wwwjdic.html>.
- Broselow, Ellen 1983. Nonobvious transfer: on predicting epenthesis errors. Language transfer in language learning, ed. by Susan Gass and Larry Selinker, 269-280. Rowley, MA: Newbury House.
- Broselow, Ellen 1992a. The structure of fricative-stop onsets. Stony Brook, NY: Stony Brook University ms.
- Broselow, Ellen 1992b. Transfer and universals in second language epenthesis. Language Transfer and Language Learning (revised edition), ed. by Susan Gass and Larry Selinker, 71-86. Amsterdam: John Benjamins.
- Browman, Catherine, and Louis Goldstein 1986. Towards an articulatory phonology. Haskins Laboratories Status Report on Speech Research 85.219-250.
- Cairns, Charles and Mark Feinstein 1982. Markedness and the theory of syllable structure. Linguistic Inquiry 13.193-225.
- Carr, Philip 2006. Universal grammar and syntax/phonology parallelisms. Lingua 116.634-656.
- Cena, Resty 1979. Double representations for some loan words in Tagalog. Studies in Philippine Linguistics 3.125-137.
- Chomsky, Noam 1964. Current issues in linguistic theory. The Hague: Mouton.
- Chomsky, Noam, and Morris Halle 1968. The sound pattern of English. Cambridge, MA: MIT Press.
- Crosbie, John. 1977. Crosbie's Dictionary of Puns. New York: Harmony Books.
- Derwing, Bruce, Maureen Dow, and Terrance Nearey 1988. Experimenting with syllable structure. Proceedings of Eastern States Conference on Linguistics (ESCOL) 5.83-94.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier, and Jacques Mehler

1999. Epenthetic vowels in Japanese: a perceptual illusion? Journal of Experimental Psychology: Human Perception and Performance 25.1568-1578.

English, Leo 1986. Tagalog-English dictionary. Manila: Congregation of the Most Holy Redeemer/National Book Store.

Ewen, Colin 1982. The internal structure of complex segments. In Hulst & Smith, 27-67.

Fleischhacker, Heidi 2002a. Cluster-dependent epenthesis asymmetries. UCLA Working Papers in Linguistics 7, Papers in Phonology 5.71-116.

Fleischhacker, Heidi 2002b. Onset transfer in reduplication. Los Angeles: UCLA ms.

Fleischhacker, Heidi 2005. Similarity in phonology: evidence from reduplication and loan adaptation. Los Angeles: UCLA dissertation.

Fowler, Carol, Rebecca Treiman, and Jennifer Gross 1993. The structure of English syllables and polysyllables. Journal of Memory and Language 32.115-140.

French, Koleen Matsuda 1988. Insights into Tagalog reduplication, infixation, and stress from nonlinear phonology. Dallas: Summer Institute of Linguistics and University of Texas at Arlington.

Fudge, Eric 1969. Syllables. Journal of Linguistics 15.253-286.

Ghani, Rayid, Rosie Jones, and Dunja Mladenec 2004. Building minority language corpora by learning to generate Web search queries. Knowledge and Information Systems 7.56-83.

Goldwater, Sharon and Mark Johnson 2003. Learning OT constraint rankings using a Maximum Entropy model. Proceedings of the Stockholm Workshop on Variation within Optimality Theory, ed. by Jennifer Spenader, Anders Eriksson, and Östen Dahl, 111-120.

- Gouskova, Maria 2003. Falling sonority onsets, loanwords, and Syllable Contact. Proceedings of the Chicago Linguistic Society 7.175-185.
- Greenberg, Joseph 1978. Some generalizations concerning initial and final consonant clusters. Universals of human language, Volume 2, ed. by Joseph Greenberg, 243-279. Stanford: Stanford University Press.
- Guest, Daniel, Gary Dell, and Jennifer Cole 2000. Violable constraints in language production: testing the transitivity assumption of Optimality Theory. Journal of Memory and Language 42.272-299.
- Hale, Mark, and Charles Reiss 2000. “Substance abuse” and “dysfunctionalism”: current trends in phonology. Linguistic Inquiry 31.157-169.
- Hall, Nancy 2003. Gestures and segments: vowel intrusion as overlap. Amherst, MA: University of Massachusetts dissertation.
- Halle, Morris 1978. Knowledge unlearned and untaught: what speakers know about the sounds of their language. Linguistic Theory and Psychological Reality, ed. by Morris Halle, Joan Bresnan, and George A. Miller, 294-303. Cambridge, MA and London: MIT Press.
- Hayes, Bruce 1999. Phonetically-driven phonology: the role of Optimality Theory and inductive grounding. Functionalism and formalism in linguistics, Volume I: General papers, ed. by Michael Darnell, Edith Moravcsik, Frederick Newmeyer, Michael Noonan, and Kathleen Wheatly, 243-285. Amsterdam: John Benjamins.
- Hayes, Bruce, Robert Kirchner, and Donca Steriade (eds.) 2004. Phonetically based phonology. Cambridge: Cambridge University Press.
- Hayes, Bruce, and Margaret MacEachern 1998. Quatrain form in English folk verse. Language 74.473–507.

- Hayes, Bruce, and Donca Steriade 2004. The phonetic basis of phonological markedness. In Hayes, Kirchner & Steriade, 1-33.
- Hayes, Bruce, and Tanya Stivers 1995. Postnasal voicing. Los Angeles: UCLA, ms.
- Hayes, Bruce, Bruce Tesar, and Kie Zuraw 2003. OTSoft 2.1. Software package, <http://www.linguistics.ucla.edu/people/hayes/otsoft/>.
- Hestenes, Magnus, and Eduard Stiefel 1952. Methods of conjugate gradients for solving linear systems. Journal of Research of the National Bureau of Standards 49.409–436.
- Hume, Elizabeth and Keith Johnson (eds.) 2001. The role of speech perception in phonology. San Diego, CA: Academic Press.
- Hura, Susan, Björn Lindblom, and Randy Diehl 1992. On the role of perception in shaping phonological assimilation rules. Language and Speech 35.59-72.
- Hulst, Harry van der and Norval Smith (eds.) 1982. The structure of phonological representations. Dordrecht: Foris.
- Hyman, Larry 2001. On the limits of phonetic determinism in phonology: \*NC revisited. In Hume & Johnson, 141-185.
- Idsardi, William (submitted). Poverty of the stimulus arguments in phonology. Ms., University of Delaware.
- Ka, Omar 1985. Syllable structure and suffixation in Wolof. Studies in the Linguistic Sciences 15.61-90.
- Karimi, Simin 1987. Farsi speakers and the initial consonant cluster in English. Interlanguage phonology: the acquisition of a second language sound system, ed. by Georgette Ioup and Steven Weinberger, 305-318. Cambridge, MA: Newbury House Publishers.

Kawahara, Shigeto (to appear). Half rhymes in Japanese rap lyrics and knowledge of similarity.

Journal of East Asian Linguistics.

Kingston, John (1990). Articulatory binding. Papers in laboratory phonology I: between the grammar and physics of speech, ed. by John Kingston and Mary Beckman, 406-434.

Cambridge: Cambridge University Press.

Kuryłowicz, Jerzy 1971. A problem of Germanic alliteration. Studies in Language and Literature in Honour of Margaret Schlauch, ed. by Mieczysław Brahmmer, Stanisław Helsztyński, and Julian Krzyżanowski, 195-201. New York: Russell and Russell.

Maclachlan, Anna E., and Mark Donohue 1999. Glottal stops and -um- in Tagalog. Abstract of paper presented at the Australian Linguistic Society Annual Conference, University of Western Australia.

McCarthy, John 2003. OT constraints are categorical. Phonology 20.75-138.

McCarthy, John, and Alan Prince 1993. Generalized Alignment. Technical Report #7, Rutgers University Center for Cognitive Science.

McCarthy, John, and Alan Prince 1995. Faithfulness and reduplicative identity. Papers in Optimality Theory; University of Massachusetts Occasional Papers, ed. by Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 18.249-384.

Minkova, Donka 2001. Testing CONTIGUITY in Middle English alliteration. Handout of paper presented at the 25<sup>th</sup> International Conference on Historical Linguistics, Melbourne.

Minkova, Donka 2003. Alliteration and sound change in Early English verse. Cambridge Studies in Linguistics 101. Cambridge: Cambridge University Press.

Murphy, Gerard 1961. Early Irish metrics. Dublin: Royal Irish Academy.

- Myers, Scott 2002. Gaps in factorial typology: The case of voicing in consonant clusters. Austin, TX: University of Texas at Austin, Ms.
- Ohala, John 1981. The listener as a source of sound change. Papers from the parasessions, Chicago Linguistic Society, 178-203.
- Ohala, John 1993. Sound change as nature's speech perception experiment. Speech Communication 13.155-161.
- Orgun, Cemil Orhan and Ronald L. Sprouse 1999. From MPARSE to CONTROL: deriving ungrammaticality. Phonology 16.191-224.
- Pater, Joseph 1999. Austronesian nasal substitution and other NC effects. The prosody morphology interface, ed. by Harry van der Hulst, René Kager, and Wim Zonneveld, 310-343. Cambridge: Cambridge University Press.
- Pater, Joseph, and Anne-Marie Tessier 2003. Phonotactic knowledge and the acquisition of alternations. In Solé, Recasens & Romero, 1777-1180.
- Pierrehumbert, Janet, and Rami Nair 1995. Word games and syllable structure. Language and Speech 38.77-114.
- Prince, Alan, and Paul Smolensky 1993/2004. Optimality Theory: constraint interaction in generative grammar. Malden, MA: Blackwell.
- Pullum, Geoffrey, and Barbara Scholz 2002. Empirical assessment of stimulus poverty arguments. Linguistic Review 19.9-50.
- Pycha, Anne, Pawel Nowak, Eurie Shin, and Ryan Shosted 2003. Phonological rule-learning and its implications for a theory of vowel harmony. West Coast Conference on Formal Linguistics 22.423-435.



- Ross, Kie 1996. Floating phonotactics: infixation and reduplication in Tagalog loanwords. Los Angeles: University of Los Angeles M.A. thesis.
- Schachter, Paul, and Fe Otanes 1972. Tagalog reference grammar. Berkeley: University of California Press.
- Selkirk, Elizabeth 1982. The syllable. In Hulst & Smith, 337-383.
- Selkirk, Elizabeth 1984. On the major class features and syllable theory. Language sound structure: studies in phonology dedicated to Morris Halle by his teacher and students, ed. by Mark Aronoff & Richard Oehrle, 1-7-113. Cambridge, MA: MIT Press. Pp. 107–113.
- Shademan, Shabnam 2002. Epenthetic vowel harmony in Farsi. Los Angeles: University of Los Angeles M.A. thesis.
- Sharp, Harold. 1984. Advertising slogans of America. Metuchen, NJ: Scarecrow Press.
- Shewchuk 1994. An introduction to the Conjugate Gradient Method without the agonizing pain. Pittsburgh: Ms., Carnegie Mellon University.
- Silverman, Daniel 1992. Multiple scansion in loanword phonology: evidence from Cantonese. Phonology 9.289-328.
- Silverman, Daniel 1997. Phrasing and recoverability. New York: Garland.
- Singh, Rajendra 1985. Prosodic adaptation in interphonology. Lingua 67.269-282.
- Solé, Maria-Josep, Daniel Recasens, and Joaquín Romero (eds.) 2003. Proceedings of the 15th International Congress on Phonetic Sciences. Barcelona: Futurgraphic.
- Steriade, Donca 1988. Reduplication and syllable transfer in Sanskrit and elsewhere. Phonology 5.73-155.

Steriade, Donca 1999. Alternatives to the syllabic interpretation of consonantal phonotactics.

Proceedings of the 1998 Linguistics and Phonetics Conference, ed. by Osamu Fujimura, Brian Joseph, and Bohumil Palek, 205-242. Prague: Karolinum Press.

Steriade, Donca 2001a. Directional asymmetries in place assimilation: a perceptual account. In Hume & Johnson, 219-250.

Steriade, Donca 2001b. The phonology of perceptibility effects: the P-map and its consequences for constraint organization. Cambridge, MA: MIT, ms.

Steriade, Donca 2003. Knowledge of perceptual similarity and its uses: evidence from half-rhymes. In Solé, Recasens & Romero, 363-366.

Steriade, Donca 2004. Syllable contact vs. syntagmatic contrast in Latin phonotactics. Paper presented at the Western Conference on Linguistics, University of Southern California.

Sulejmenova, B.A. 1965. O fonetičeskom osvoenii leksiki, zaimstvovannoj iz russkogo jazyka v kazaxskij. Progressivnoe vlijanie russkogo jazyka na kazaxskij, ed. by S.K. Kenesbaev, V.A. Isengaliev, Sh.Sh. Sarybaev, and S. Nurkhanov, 60-95. Alma-Ata: Nauka.

Treiman, Rebecca 1983. The structure of spoken syllables: evidence from novel word games. Cognition 15.49-74.

Urdang, Laurence and Celia Robbins (eds.) 1984. Slogans. Detroit: Gale Research Company.

Vennemann, Theo 1988. Preference laws for syllable structure and the explanation of sound change: with special reference to German, Germanic, Italian, and Latin. Berlin: Mouton de Gruyter.

Warner, Natasha, Allard Jongman, Anne Cutler, and Doris Mücke 2001. The phonological status of Dutch epenthetic schwa. Phonology 18.387-420.

- Weide, Richard (ed.) 1995. Carnegie-Mellon University Pronouncing Dictionary. Available at [www.speech.cs.cmu.edu/cgi-bin/cmudict](http://www.speech.cs.cmu.edu/cgi-bin/cmudict).
- Weijer, Jeroen van de 1996. Segmental structure and complex segments. Tübingen: Max Niemeyer.
- Wilson, Colin 2003. Experimental investigation of phonological naturalness: consonant harmony vs. random alternation. West Coast Conference on Formal Linguistics 22.533-546.
- Wilson, Colin 2006. Learning phonology with substantive bias: an experimental and computational study of velar palatalization. Cognitive Science 30.945-982.
- Wright, Joseph 1910/1954. Grammar of the Gothic Language. 2<sup>nd</sup> edition 1954, 1966 reprinting. London: Oxford University Press.
- Wright, Richard 1996. Consonant clusters and cue preservation in Tsou. Los Angeles: UCLA dissertation.
- Yip, Moira 1993. Cantonese loanword phonology and Optimality Theory. Journal of East Asian Linguistics 2.261-291
- Yu, Alan 2003. The morphology and phonology of infixation. Berkeley, CA: UC Berkeley dissertation.
- Yu, Alan 2004. Explaining final obstruent voicing in Lezgian: phonetics and history. Language 80.73-97
- Zhang, Jie 2001. The contrast-specificity of positional prominence: evidence from diphthong distribution. Paper presented at the annual meeting of the Linguistic Society of America, Washington, DC.
- Zhang, Jie and Yuwen Lai (submitted). Testing the role of phonetic naturalness in Mandarin tone sandhi. Lawrence, KS: University of Kansas, ms.

Zhang, Jie, Yuwen Lai, and Craig Turnbull-Sailor (in progress). Revisiting the psychological reality of Taiwanese tone sandhi. Lawrence, KS: University of Kansas, ms.

Zsiga, Elizabeth, Maria Gouskova, and One Tlale 2006. On the status of voiced stops in Tswana: Against \*ND. Proceedings of the Northeast Linguistic Society 36.

Zuraw, Kie 2005. Markedness in the distribution of an optional rule. Los Angeles: UCLA, ms.

---

<sup>1</sup> Even if the scope of linguistic inquiry is only language-specific cognitive mechanisms, we must still understand domain-general mechanisms' influence on linguistic behavior, if to factor it out.

<sup>2</sup> See Hura et al. (1992) and discussion in Steriade (2001a) however: misperceptions in this environment are mostly nonassimilatory. This example is chosen for its conceptual clarity, but there may be other examples in which the phonetic data are more straightforward.

It would also be possible for the language-transmission explanation to include an element of variation in pronunciation by adults (see Blevins 2004), for example a bias towards mispronouncing /np/ as [mp] but not /pn/ as [pm], in this example. But then we must address the question of whether such variation is itself governed by mental biases or could be purely mechanical in the vocal tract.

<sup>3</sup> See also earlier work on 'crazy rules' (Bach & Harms 1972) and the unnaturalness of phonology (Anderson 1981).

<sup>4</sup> Carr 2006 states 'there is no poverty of the stimulus argument in phonology', because '[p]honological objects and relations are internalisable [i.e. available in the speech signal]' (p. 654). Carr is contrasting the relation he takes to be important in phonology—sequential order—to the more abstract, hierarchical relations necessary for describing syntax. I don't think Carr is arguing that a poverty-of-the-stimulus argument can never be made in phonology—that is, that speakers can never be shown to have phonological knowledge that is unavailable in the learning data.

<sup>5</sup> As noted later in this section, it is also possible that borrowers merely misperceive the source word in the first place.

---

<sup>6</sup> Why a difference between m and n in this apparently sonority-based scale? It can be argued that [n] is more vowel-like than [m] because nasal-antiformants that might interfere with vowel-like formant structure are higher (and thus interfere less) for [n] than for [m]. See Zuraw (2005) for a discussion of this, based on an idea of Daniel Silverman.

<sup>7</sup> Karimi (1987) documents prothesis for sT, sm, sn, sl, and anaptyxis for TL (all in agreement with Fleischhacker), but does not investigate SR.

<sup>8</sup> Though see fn. 5. In any case, sr is treated as more splittable than Tr.

<sup>9</sup> Although Gothic has other initial clusters besides fr, gr, sl, st, sk, they appear not to be attested with reduplication.

<sup>10</sup> The 1,964 puns in Fleischhacker's corpus come from a book of puns (Crosbie 1977), two books of product slogans (Sharp 1984, Urdang and Roberts 1984), and assorted media sources.

<sup>11</sup> Napoleon Blown-aparte: title of a 1966 cartoon in the "Inspector" series, referring to a mad-bomber character (www.imdb.com). The pun, which probably predates the movie, consists in juxtaposing the explicit blown-apart(e) with the implied Bonaparte (the name that usually follows Napoleon).

<sup>12</sup> For puns of the form  $C_1C_2V... \sim C_1VC_2V...$ , such as broke ~ baroque, there are not enough tokens to draw conclusions about cluster differences (though the trend is in the predicted direction, with relatively many TL clusters and relatively few ST). Only one pun of the form  $C_1C_2V... \sim VC_1C_2V...$  (steamed ~ esteemed) occurs.

<sup>13</sup> As in the case of puns, it is unclear which of these differences are significant. Minkova gives Middle-English dictionary counts for each initial cluster, so it is straightforward to determine whether a given cluster alliterates cohesively at a higher than chance level (most do). But,

---

determining whether two higher-than-chance rates of cohesive alliteration are significantly different probably requires a Monte Carlo simulation.

<sup>14</sup> In the case of loan adaptation, the preference plays out only in those speakers who have access to the foreign source form.

<sup>15</sup> See Fleischhacker for discussion of  $C_1C_2\alpha \sim C_1\emptyset C_2\alpha$  pairs, with schwa epenthesis.

<sup>16</sup> Fleischhacker 2002a and Andersen 1972 (p. 36) point out that modifications of this type are common in casual, emphatic English (e.g. *puh-leeze* ‘please’, *kee-rist* ‘Christ’).

<sup>17</sup> Fleischhacker also reports a second rating experiment, focusing on S-nasal, S-liquid, SW. The trend seems to be for word/split-word pairs to be rated higher for SW than for S-liquid and higher for S-liquid than for S-nasal, as expected, but the trend is very slight and Fleischhacker does not report on its statistical significance.

<sup>18</sup> Unless enclosed in square brackets, all examples are given in normal Tagalog spelling, with the possible addition of hyphens, boldface, and italics. Examples in square brackets are phonetic transcriptions.

<sup>19</sup> There is also a rarer pattern, gumaraduate, pinorotekta-han; see section 7.1 for some discussion of epenthetic vowels, found especially in older loans.

<sup>20</sup> Magbabalaod and magbabalaud are probably intrusions from Cebuano. Like the CorpusBuilder software, this method has difficulty keeping out text from other Philippine languages.

<sup>21</sup> There are some loans beginning in nasal-glide or liquid-glide (mw, my, ny, ly), but no infixes were found in the corpus. There are also loans beginning in fl or fr that take infixes, but none beginning in fw or fy (that take infixes) to compare them to.

---

<sup>22</sup> Since these are not true type frequencies but token-weighted type frequencies, it is unclear whether Fisher's Exact test might be overly sensitive, insufficiently sensitive, or just right here. All counts were rounded to the nearest integer in order to apply Fisher's Exact Test.

A multifactor ANOVA was also performed on these data, with each word a trial, percent split as the dependent variable, and, as factors,  $C_1$ ,  $C_2$  (all stops combined), infix (um or in), etymology, other morphology (e.g. prefix i and suffix an), shape of stem's first syllable (open or closed), stress of stem's first syllable, and reduplication (yes or no). Cells were unbalanced, with many empty accidentally or for systematic reasons. Eliminating factors without significant main effect and not participating in significant interactions,  $C_2$  has a significant main effect ( $F(3,159)=10.56$ ,  $p<.0001$ ) and participates in no significant interactions. The significant pairwise differences ( $p<.05$  by Tukey's HSD) are  $C_2=l$  vs.  $C_2=w$ ,  $C_2=r$  vs.  $C_2=w$ , and  $C_2=r$  vs.  $C_2=y$ . The other strong main effect is of reduplication (which participates in no significant interactions),  $F(1,161)=46.84$ ,  $p<.0001$ : reduplicated words are less likely to undergo splitting (see section 7.1 for examples of reduplicated, infixed words). Whether the first syllable of the stem is stressed has a significant main effect,  $F(1,161)=4.82$ ,  $p<.01$ , and participates in no interactions: splitting is more likely when the stem's first syllable is stressed. This seems to be in line with Avery and Lamontagne 1995, though they describe data that involve infixation with epenthesis. Finally,  $C_1$  has significant interactions with infix type (um or in) and other morphology, as well as a significant main effect, but given the small number of items in each cell, I have not attempted to dissect these effects.

<sup>23</sup> Thanks to participants in the UC Berkeley linguistics colloquium for pointing this out.

<sup>24</sup> It is not always easy to determine whether a word is a Spanish loan. Translado 'translated' for example, looks Spanish, but is not a real word in Spanish (where 'translated' is traducido and



‘moved’ is traslado). More likely, it is the English word translate altered to look more Spanish—and thus more Tagalog, since Spanish loans have been in the language much longer and are better incorporated—by using the English-to-Spanish ated/ado correspondence. Other alterations are not so easy to detect. For example, is transporma from Spanish transforma (conjugated form of transformar ‘to transform’) or from English transform, with the a added to give a more Spanish appearance? Clearly English-origin items such as traslado were excluded from the Spanish-origin counts, but ambiguous cases such as transporma were included.

<sup>25</sup> The team name is Eskumor; this form is based on ‘score’ with, unusually, prothesis but an infix after the cluster. Also unusual is the prothetic vowel e rather than i. ‘Score’ is usually adapted as iskor, with infixed um-iskor. From the team’s website, at [eskumor.sitesled.com/about.html](http://eskumor.sitesled.com/about.html): ‘Bonn Reyes invented the name “Eskumor” after mistakenly pronouncing the word “umiskor” to “iskumor” or “eskumor”, resulting in a team huddle chant for six years. In 2002, it became the new monicker of the Bloomfield Basketball team before it was disbanded in 2006.’

The other tokens are scrinutimize, iskinetch (from sketch—this word may have the prefix i or be formed similarly to eskumor), slinice, sinlow, sprinayan (from spray, with the suffix -an), spinray-paint, stinalk and stino-stalk, strumay, and struming.

<sup>26</sup>Unprothesized sC-initial clusters are somewhat more common with reduplication than with infixation (e.g. pag-sno-snorkel ‘snorkeling’, mag-si-sleep ‘will sleep’)—278 tokens were found. Still, there are too few attested types for each cluster category to get a sense of whether the expected reduplication pattern is followed, with s-stop simplified the least often and s-glide the most often.

---

<sup>27</sup> The full ANOVA is applied only to subjects who rated all 6 cluster types. Pairwise comparisons include a few more subjects.

<sup>28</sup> Ross 1996 attempts to repair the NOCODA analysis by adding variably ranked \*COMPLEX, which would prefer g-um.-rad.wet. If, however, \*COMPLEX stands for a family of constraints requiring a consonant to be adjacent to segments that allow expression of its acoustic cues (Steriade 1999), this makes incorrect predictions about which clusters should split more often. See the discussion of cluster markedness in section 8.2. Moreover, language-internal evidence requires that \*COMPLEX >> NOCODA, since word-internal clusters are syllabified heterosyllabically (ak.lat ‘book’).

It might be objected that LEFT-ANCHOR is violated in vowel-initial words such as abot, ‘infixed’ as um-abot ‘attain’. But, words spelled (and often transcribed) with an initial vowel actually begin with a glottal stop (unless preceded by a consonant-final word within the same phrase, in which case the glottal stop is optional). If this glottal stop is underlying, then the infix form ?-um-abot does satisfy LEFT-ANCHOR. If the glottal stop is epenthetic, then the constraints requiring its insertion force LEFT-ANCHOR to be violated no matter what (the word cannot begin with a), so LEFTMOST pushes the infix as far to the left as possible.

<sup>29</sup> A question not addressed here is why an infix can’t move to the most splittable site in the stem. If CONTIG-Ca/u >> CONTIG-Cu/u, we expect labusaw ‘made turbid’, to be infixated as \*lab-um-usaw (actual form l-um-abusaw ‘to make turbid’). We can rule out \*lab-um-usaw in Tagalog with a categorical alignment constraint (McCarthy 2003) forbidding um from occurring later than the first syllable, but the problem remains on a typological level: why do no languages behave that way? Similar typological problems arise in all standard-OT approaches to infixation: if the constraints on infix placement are freely rankable with other markedness constraints, we

---

predict languages in which infixes can travel wherever needed to repair markedness violations, such as to the sites of bad syllable contacts (pad-um-nara).

<sup>30</sup> A fuller model would derive the ranking values from perceptual similarity. See Wilson (2006) for a model that derives faithfulness constraint weights from perceptual confusion data.

<sup>31</sup> Re ‘ordinary methods’: as Idsardi (submitted) points out, there is no such thing as a purely data-driven learner—we can’t contrast learners with no expectations or biases to learners with some. We can only contrast learners with different sets of expectations, such as a learner with various domain-general expectations and a learner with those plus some language-specific expectations.

<sup>32</sup> Shelley Velleman (p.c.) raises the possibility that, if the TR-TW difference were already in place (perhaps because of epenthesis at an earlier stage), speakers could pick up on sonority as an important factor in determining splittability and extend that factor’s applicability to the SC cases. This would require implicit knowledge of sonority differences, but the bias about how to apply those differences would come from overt evidence.

<sup>33</sup> Cena (1979) assumes that splitting of a loan cluster by the infix (and partial reduplication) results from an extra vowel, but in the examples he considers the vowel is robust (and spelled).

<sup>34</sup> Many loans that, in the source language, begin consonant-glide can optionally be spelled with an extra vowel in Tagalog: byahe, biyahe ‘travel’, from Spanish viaje. In the corpus data, only tokens spelled without this extra vowel were used. It is possible that sometimes the extra vowel is pronounced though not spelled. The reverse does seem to occur, as attested by reduplicated forms in the corpus such as ba-biyahe. The vowel a in the reduplicant makes sense only if the stem is treated as bya.he, not bi.ya.he.

<sup>35</sup> Thanks to Colin Wilson and Christian Uffmann for raising this possibility.

---

<sup>36</sup> At a reviewer's suggestion, this idea was also implemented using not observed frequencies of the clusters, but ratios of observed to expected frequencies (based on the independent stem-initial frequency of the second member of each cluster). Taking expected values from a database of disyllabic native roots of Tagalog (drawn from English 1986's dictionary) or from the Carnegie-Mellon Pronouncing Dictionary of English (Weide 1995), results were similar: sT and sn had high O/E values, whereas sm, sl, shr, and sw had low O/E.

<sup>37</sup> Generalizations of type, 'insert the infix after the first X' have little hope of working, because most of the generalizations would fare so badly on the native data: X=[-syllabic] does well, but X=[-continuant], for example, which must receive a large weight in order to favor st-in-op over s-in-top, is falsified by abundantly many forms such as l-um-akad, not \*lak-um-ad.

<sup>38</sup> I'm greatly indebted to Colin Wilson for sharing his software that implements Conjugate-Gradient learning of MaxEnt weights, and then generation using those weights, and for making the adjustments necessary to allow the software to run on my system.

<sup>39</sup> Training frequencies for sonorant Cs were taken directly from the corpus (with the exception of l and h with um). Comparing these counts to the number of roots in the root database starting with each sonorant consonant, a ratio of corpus occurrence to root-database occurrence was then obtained for each infix. To avoid excessive hand-checking, counts for obstruent Cs, for vowel/glottal-stop initial roots, and for l and h with um (where there is no variation, only infixation) were simulated by counting, for example, the number of ba... roots in the root database, and multiplying this number by the corpus/root-database ratio. Stem beginning with n and taking the in/ni affix are difficult to classify: is ninakaw 'be robbed' n-in-akaw or ni-nakaw? There were only 17 such words in the corpus (nasal-initial roots are underrepresented in Tagalog), and each was counted as half infixated and half prefixed.

---

<sup>40</sup> Warner et al. (2001) argue that the schwa in forms like [kələm] results from a separate vocalic gesture, because the articulation of [l] in forms like [kələm] patterns more with [l] before underlying schwa than with [l] in forms like [kəlm]. Hall counters that the articulatory difference between the [l] articulations in [kələm] and [kəlm] could result from the timing difference, rather than from a true epenthesis.

<sup>41</sup> Release and target are terms referring to landmarks within a gesture (Browman & Goldstein 1986). In temporal order, the gestural landmarks are onset, target, center, release, and offset. If the release of C<sub>1</sub> coincides with the target of C<sub>2</sub>, there is no interval of open vocal tract between the two consonants.

<sup>42</sup> This is not exactly faithful to Hall's account of svarabhakti vowels. She proposes a general constraint, applying to all consonants, requiring alignment of C<sub>1</sub>'s release to C<sub>2</sub>'s target, and a specific constraint for obstruent-sonorant clusters requiring obstruent C<sub>1</sub>'s center to be aligned with sonorant C<sub>2</sub>'s onset, a configuration that results in an excrescent vowel. These two constraints would both be violated by infixation into an obstruent-sonorant cluster.

<sup>43</sup> As mentioned in fn. 28, this argues against using \*COMPLEX to explain the existence of infixation variants in which the infix splits the onset cluster: if \*COMPLEX is viewed as a complex of constraints against complex onsets of varying degrees of markedness, then the wrong prediction is made about which clusters should split most easily.

<sup>44</sup> Steriade 2004, however, proposes that in Latin, CW clusters are more marked than other clusters.

<sup>45</sup> The main reason to believe that the vowel is deleted, not inserted, is that native Tagalog lexical roots obey a disyllabic minimum. It would be an odd coincidence if all the underlyingly

---

monosyllabic native roots began with consonant-glide clusters (and almost no disyllabic or longer roots began with such clusters).

<sup>46</sup> Counts are from disyllabic native roots found in English 1986.



	ST	Sm	Sn	Sl	SR	SW
Sinhalese	VST				SVR	
Farsi	VST	VSm	VS <sub>n</sub> <sup>1</sup>	VSl	SVR <sup>2</sup>	sVw/sVv <sup>3</sup>
Hindi (as described by Bharati)	VST	VSm	VS <sub>n</sub> ~SV <sub>n</sub>	VSl~SVl	SVR <sup>4</sup>	
Wolof	VST	VSm~SVm	VS <sub>n</sub> ~SV <sub>n</sub>	VSl~SVl	? <sup>5</sup>	VS <sub>w</sub> ~SV <sub>w</sub>
Kazakh	VST	VSm~SVm	VS <sub>n</sub> ~SV <sub>n</sub>	SVl		
Hindi (as described by Singh/Broselow)	VST			SVl		
Egyptian Arabic	VST			SVl		SVW
Japanese	SVT	SVm	SV <sub>n</sub>	SVr	SVR	SV <sub>w</sub> <sup>6</sup>

Table 1

Cluster adaptation patterns in languages that allow no word-initial CC clusters.

Cells with anaptyxis are shaded lightly; cells with prothesis are shaded darkly; cells with variation have intermediate shading.

<sup>1</sup> When asked to adapt visually presented words into Farsi, Shademan's (2002) four subjects agreed on esnupi for 'Snoopy', following the pattern for established loans, but produced fenabel for novel 'Schnabel'. Karimi (1987) reports [e]snow for 'snow' in the English speech of her Farsi-speaking consultants.

<sup>2</sup> Karimi (1987) and Shademan (2002) both state that prothesis occurs for all SC clusters, do not investigate SR. Fleischhacker's data on fr and sw come from Shademan herself.

<sup>3</sup> A third pattern is vocalization of w, as in [su?et] 'sweat' (Karimi 1987: 311).



---

<sup>4</sup> Bharati states that Sr is usually left intact (i.e. not nativized at all), but that if the cluster does undergo epenthesis, the epenthesized form given is SVr.

<sup>5</sup> Fleischhacker states that the speaker she consulted displayed variation for sibilant-liquid clusters, as in [eslɛpnir] ‘Sleipnir’ but [solovaki] ‘Slovakia’. Only one example is given for Sr, however, [siri laŋka] ‘Sri Lanka’.

<sup>6</sup> Japanese is substituted here for Fleischhacker’s Korean. There is some variation for sw items in Japanese: e.g. suwahiri ‘Swahili’, but suetto ‘sweat’ (and, a much rarer pattern, seetaa ‘sweater’). (Data from Breen, n.d.)



reduplication (Fleischhacker)	ST		>			T-liquid			
puns (Fleischhacker)	ST	>	S-liquid	>		T-liquid			
alliteration (Fleischhacker, Minkova)	ST	>	SN	>	Sl	>	SR	,	T-liquid
discrimination experiment (Fleischhacker)	ST	,	S-sonorant	>		T-sonorant			
similarity-rating experiment (Fleischhacker)	ST	>	SN	>	S-liquid	,	SW	,	T-liquid
infixation game (Pierrehumbert & Nair)	ST	>	S-liquid	,		T-liquid			

Table 2

Summary of cluster distinctions discussed in section 2.1

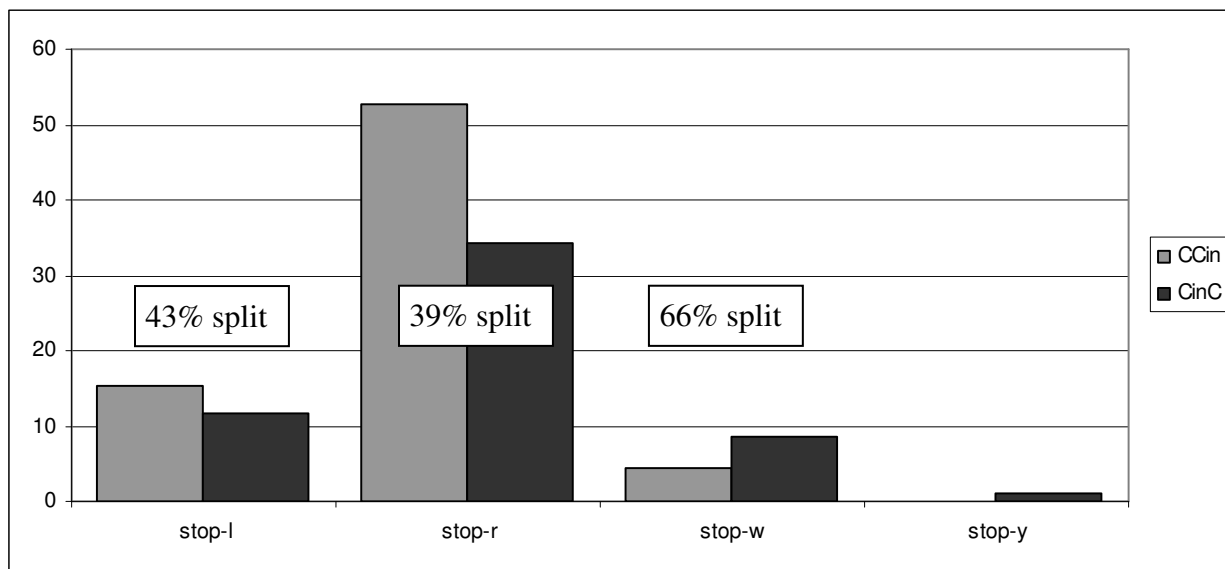


Figure 1

Token-weighted type frequencies for splitting vs. nonsplitting: all loans in corpus with in

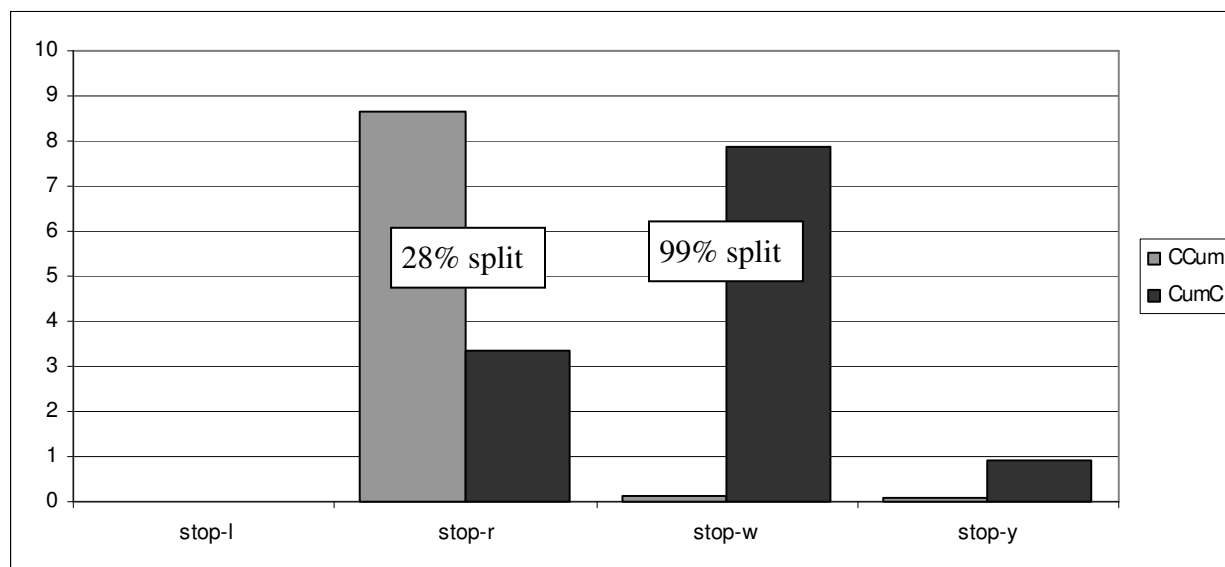


Figure 2

Token-weighted type frequencies for splitting vs. nonsplitting: all loans in corpus with um

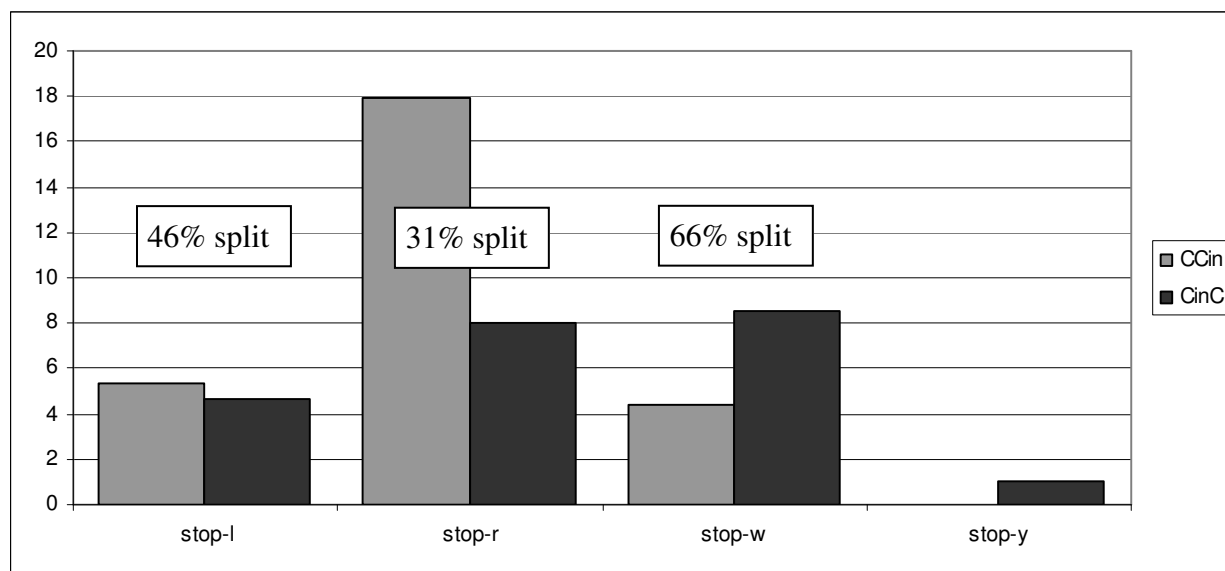


Figure 3

Token-weighted type frequencies, splitting vs. nonsplitting: Spanish-etymology loans only, infix

in

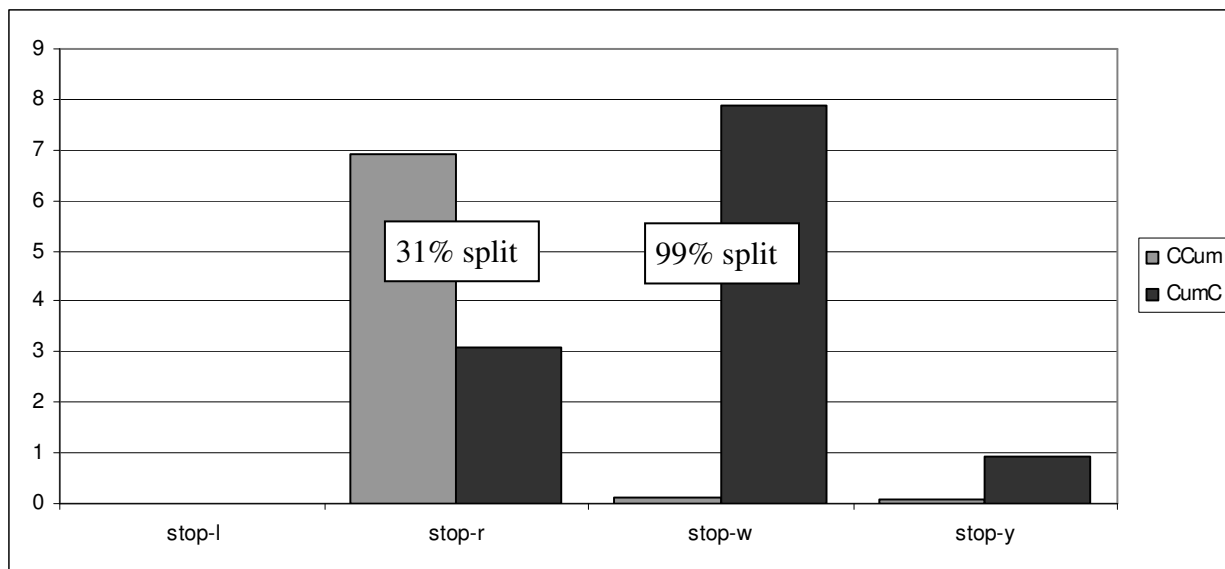


Figure 4

Token-weighted type frequencies, splitting vs. nonsplitting: Spanish-etymology loans only, infix

um

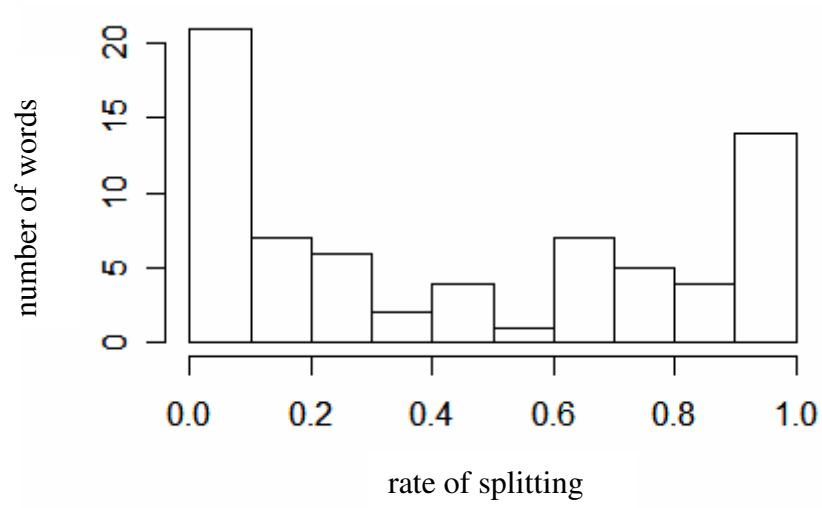


Figure 5

Histogram of splitting rates for CC-initial words, minimum frequency 5



**Alam ba ninyo kung ano mayroon ang mga salitang ito: akala, asal, asam, kubol, hukom, halal, hamak, hikayat?**

Lahat sila ay galing sa wikang Arabe. Ang mga salitang ito ay dumating sa Pilipinas sa pamamagitan ng Indonesia at Malaysia.

**Ano ang pinakakaraniwang pinakamahabang (kung titik ang pag-uusapan) ginagamit na salita sa Tagalog na hindi hiniram sa ibang wika?**

Buuin ang mga sumusunod na 2 halimbawa para malaman ang sagot.

*Piliin ang salita na pinakamalamang na pupuno sa patlang:*

Anong pelikula ay \_\_\_\_\_ sa Baler nung 1976?

syinuting

sinyuting

*Markahan ninyo ang bawat pagpipilian mula sa antas na 1 hanggang 7.*

	di-pinakamalamang	1	2	3	4	5	6	pinakamalamang
syinuting		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	di-pinakamalamang	1	2	3	4	5	6	pinakamalamang
sinyuting		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Pindutin ang boton na Kasunod, kapag handa kayo nang magpatuloy.

Ito ay ang pangungusap 3 sa loob ng 14 pangungusap.

Figure 6

Sample page of survey: forced-choice task and ratings task

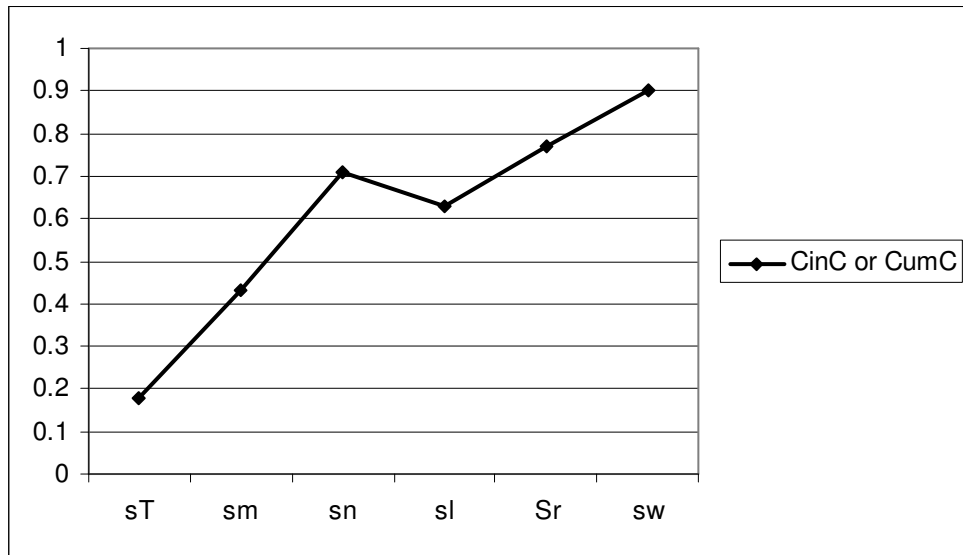


Figure 7

Results of forced-choice task: rate at which split option was chosen, for each cluster

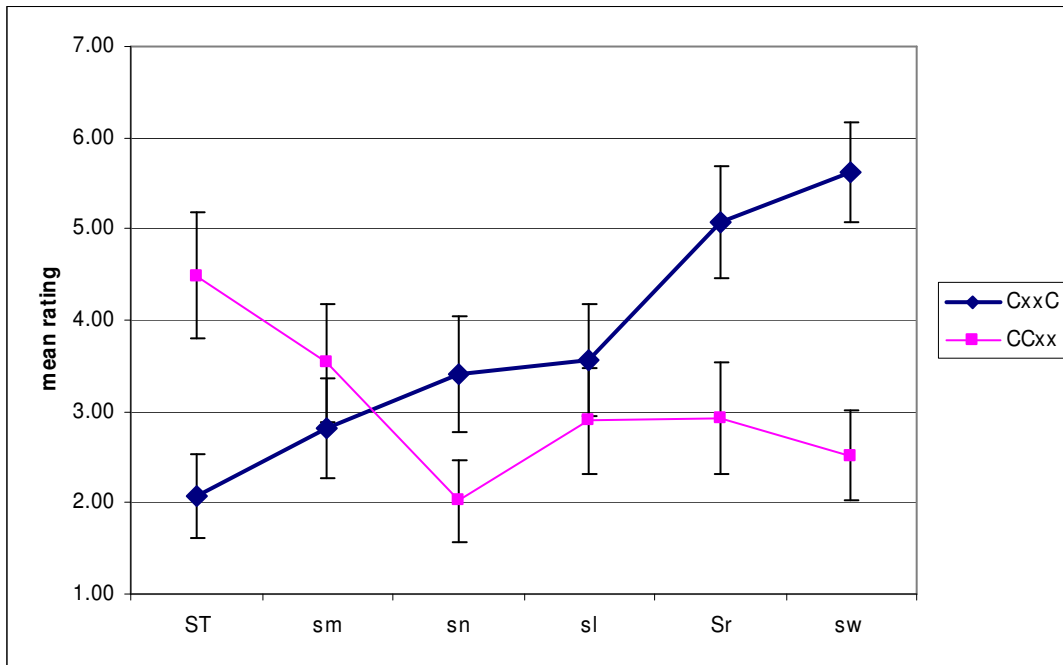


Figure 8

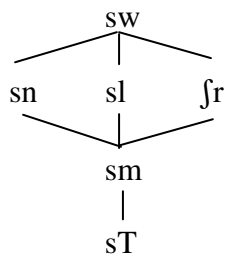
Results of ratings task: mean ratings for both options, for each cluster. Error bars indicate 95% confidence interval.

		<u>sT</u>	<u>sm</u>	<u>sn</u>	<u>sl</u>	<u>sr</u>
<u>sm</u>	rating differences	p=.0892				
	choices	p=.0049				
<u>sn</u>	rating differences	p<.0001	p=.0019			
	choices	p<.0001	p=.0045			
<u>sl</u>	rating differences	p=.0021	p=.0031	p=.7305		
	choices	p<.0001	p=.0344	p=.8392		
<u>Sr</u>	rating differences	p=.0003	p=.0003	p=.1539	p=.0666	
	choices	p<.0001	p=.0005	p=.2859	p=.0945	
<u>sw</u>	rating differences	p<.0001	p<.0001	p=.0030	p=.0021	p=.0838
	choices	p<.0001	p<.0001	p=.0148	p=.0018	p=.0953

Table 3

Significance of pairwise differences between clusters in survey results. Cells are shaded when  $p < .05$  for at least one measure.

most splittable



least splittable

Figure 9

Significant differences in splittability, from survey data

112.000	ANCHOR-STEM
99.387	CONTIG-ST/V
97.543	CONTIG-Sm/V
97.355	LEFTMOST
97.075	CONTIG-Sn/V
96.398	CONTIG-Sl/V
95.206	CONTIG-Sr/V
93.036	CONTIG-SW/V

Table 4

Boersmian ranking values<sup>1</sup>


---

<sup>1</sup> In Boersma's system, a constraint ranking is created for each instance of generation: each ranking value is perturbed somewhat by the addition of a random variable, and the resulting numbers are used to order the constraints (thus, a constraint with a higher ranking value has a tendency to be ranked higher). The constraint ranking thus derived chooses an output candidate in standard OT fashion. Over many iterations, the frequency of an output candidate is in proportion to the total probability of the rankings that derive it. In Table 4, for example, CONTIG-ST/V is fairly likely to outrank LEFTMOST (so splitting of ST results a bit over 20% of the time, and shown in Figure 10) while CONTIG-Sm/V is only somewhat likely to outrank LEFTMOST (producing splitting of sm almost half the time), and CONTIG-Sn/V is somewhat likely to be ranked below LEFTMOST (producing splitting of sn a bit more than half the time), etc.

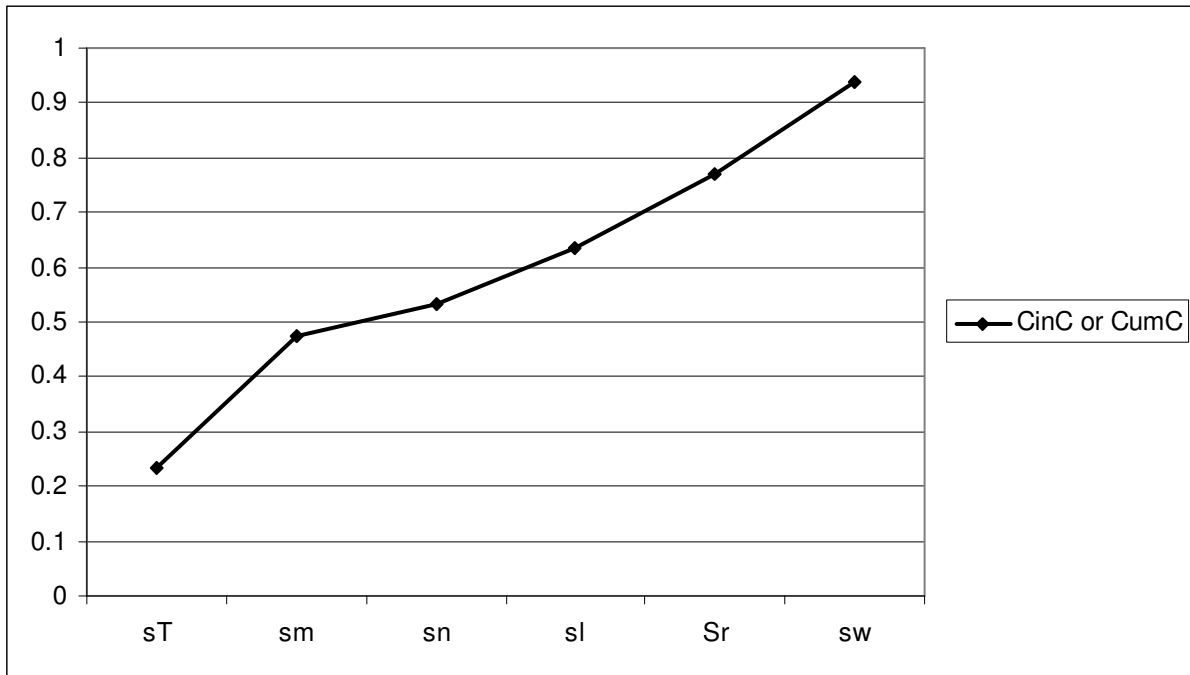


Figure 10

Splitting rates generated by grammar in Table 4

I.	II.	III.	IV.	
Onset copied, not split by infix	Onset copied, split by infix	Onset simplified, C <sub>2</sub> skipped	Onset simplified, C <sub>2</sub> vocalized (if C <sub>2</sub> is glide)	
	0 g-um-wa-gwapo	1 g-um-a-gwapo	12	0 ‘be handsome’
	0 s-um-we-sweldo	3 s-um-e-sweldo s-um-i-sweldo	7 s-um-u-sweldo	33 ‘pay salary’
kw-in-e-kwenta	1	0 k-in-e-kwenta	2 k-in-u-kwenta	20 ‘count’
	0 b-um-ya-byahe	3 b-um-a-byahe	22 b-um-i-byahe	4 ‘travel’
pr-in-o-problema, pr-in-u-problema	28 p-in-ro-problema p-in-ru-problema	3 p-in-o-problema p-in-u-problema	249 N.A.	‘have problem’
pr-in-o-promote, pr-in-u-promote	11 p-in-ro-promote p-in-ru-promote	1 p-in-o-promote p-in-u-promote	54 N.A.	‘promote’
	0 p-in-re-prepare	1 p-in-e-prepare p-in-i-prepare p-in-e-prepara	10 N.A.	‘prepare’
	0 p-in-ri-prito	2 p-in-i-prito	32 N.A.	‘fry’

Table 5

Corpus-attested variants for reduplication+infixation, with token frequencies



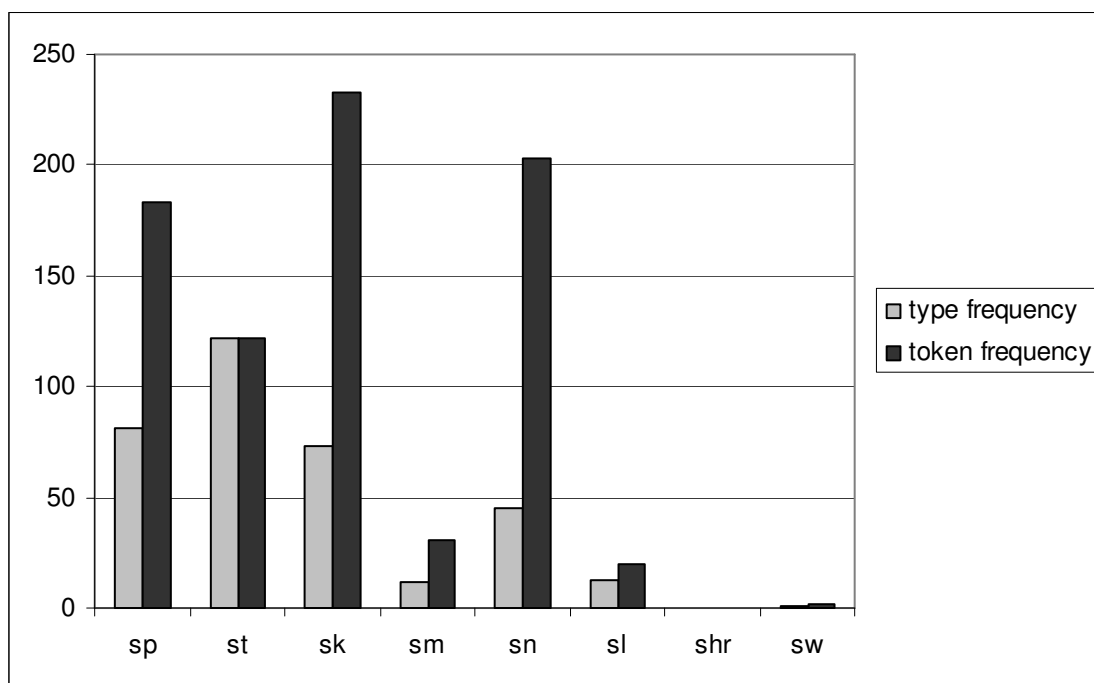


Figure 11

Frequency with which English SC-initial loans appear with a prothetic vowel

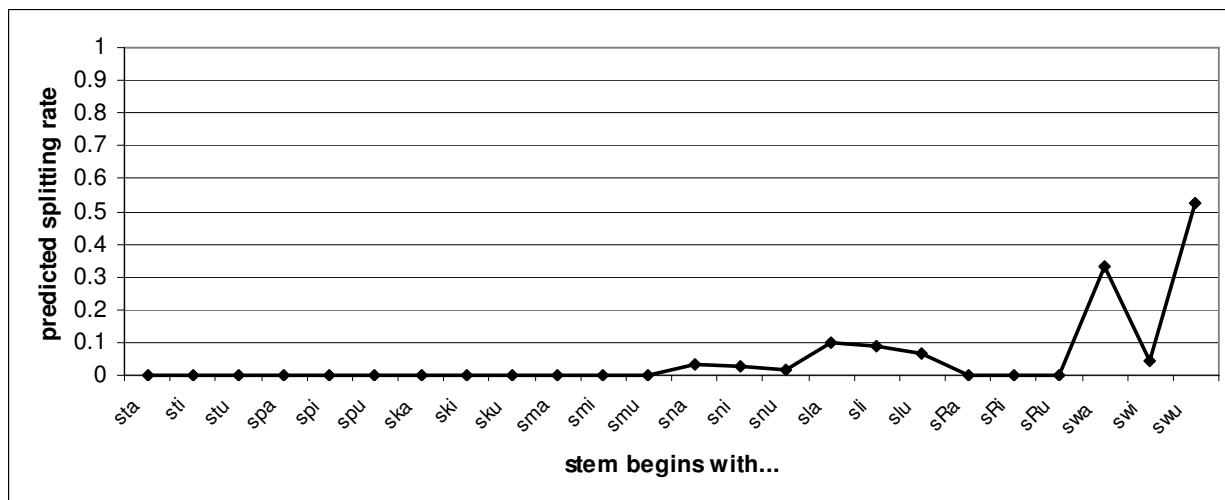


Figure 12

Splitting rates predicted by MaxEnt model

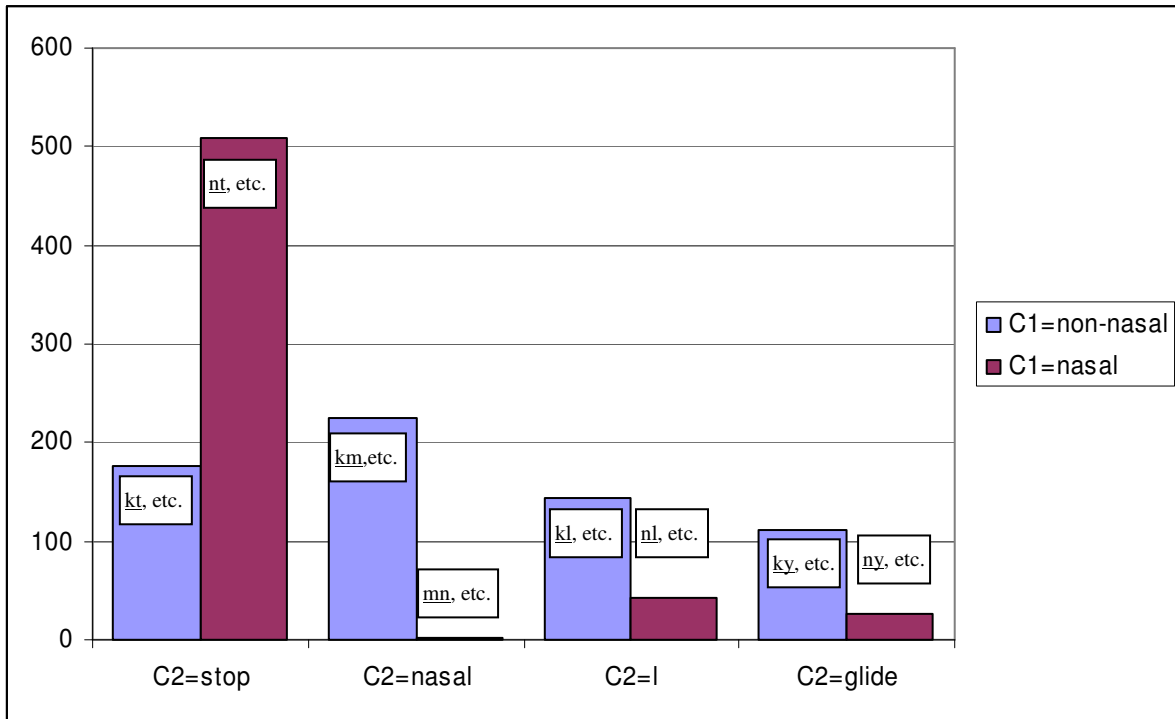


Figure 13

Type frequencies in dictionary of root-internal CC clusters