THE ROLE OF PHONETIC KNOWLEDGE IN PHONOLOGICAL PATTERNING: CORPUS AND SURVEY EVIDENCE FROM TAGALOG INFIXATION

Kie Zuraw

University of California, Los Angeles

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 article 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 Fleischhacker’s crosslinguistic findings of cluster splittability. This article argues that explaining the data requires attributing to Tagalog speakers phonetic knowledge and a bias about how to apply that knowledge.*

1. INTRODUCTION. Generative linguistics seeks to describe the mental apparatus (language-specific and otherwise) 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 Hura et al. 1992, Steriade 2001a), 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 in

* Many people besides the author have put substantial work into this article. For detailed critiques of every aspect of the article, I’m indebted to Brian Joseph, Jaye Padgett, Donca Steriade, and two anonymous referees. 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 article were programming work by Ivan Tam; a grant from the UCLA Faculty Senate; and earlier work by Rayid Ghani, Rosie Jones, and Dunja Mladenic, 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.

1 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 only to factor it out.
Steriade 2001a, for example, and it 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 & 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 mishear /an + pa/ as [ampa], but to correctly hear /an + i/ as [ani]—that is, to mishear 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) and Blevins (2004) within the framework of evolutionary phonology. See also Ohala 1981, 1993, Hale & 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, and 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. For example, under the classic optimality theory (OT) idea that the constraint set is universal (Prince & Smolensky 2004 [1993]), one 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.

---

2 But see Hura et al. 1992 and discussion in Steriade 2001a: 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 toward mispronouncing /np/ as [mp] but not /pn/ as [pm], in this example. But then one must address the question of whether such variation is itself governed by mental biases or could be purely mechanical in the vocal tract.

3 See also earlier work on ‘crazy rules’ (Bach & Harms 1972) and the unnaturalness of phonology (Anderson 1981).
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 2007 and Zhang et al. 2007). This type of research has included artificial language-learning experiments (Guest et al. 2000, Pater & Tessier 2003, Pycha et al. 2003, Wilson 2003, 2006), including novel language games (Treiman 1983, Derwing et al. 1988, Pierrehumbert & Nair 1995), and the study of second-language phonology (Broselow 1992a,b). Less commonly, there has been research on literary invention, such as puns, rhymes, and alliteration, mostly using corpora (Minkova 2001, 2003, Fleischhacker 2001b, 2005, Steriade 2003, Kawahara 2007). 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 et al. 1999) and uncertainty about 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, and so forth. 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 [baxsz]? 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. I argue that the Tagalog evidence supports the existence of a mental bias.

As in most of the works just cited, the structure of my argument is along the same lines as Pullum and Scholz’s (2002) definition of argument from poverty of the stimulus (see §6). That is, I argue that speakers 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 perceptual-similarity account (2001a,b, 2005), and explain the relevance of Tagalog infixation. I then present evidence from a written

---

4 Carr (2006:654) 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]’. 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.
corpus of Tagalog and from a survey of Tagalog speakers. I argue 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, and finally consider alternative explanations of the data.

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 is that stop-sonorant clusters (TL) are more splittable by an epenthetic vowel than are sibilant-stop clusters (ST) (Broselow 1983, 1992a,b, Singh 1985, Fleischhacker 2001a). The pattern found in Farsi (from Fleischhacker 2001a; 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). 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 & Goldstein 1986). See §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 an ST-initial word produces the unmarked syllable contact S.T (V.S.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.

Fleischhacker’s explanation (2001a,b, 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,5 and that TL and TVL are more similar to each other than are SI 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

5 As noted later in this section, it is also possible that borrowers merely misperceive the source word in the first place.
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, by contrast, 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.

In the remainder of this section I summarize 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 C1 (stop vs. sibilant) and C2 (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 C1 relatively constant—in the examples below, mostly [s] with some [ʃ] and [z]—and vary C2. This is what Fleischhacker (2001a, 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 < Sm < Sn6 < Sl < SR, SW

\[S = \text{ sibilant, } T = \text{ stop, } R = \text{ rhotic, } W = \text{ glide}\]

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 cutoff after Sm is exemplified by Hindi (as described in Bharati 1994: 56–59), with prothesis for ST and Sm, and variation or anaptyxis for the rest. Kazakh (Sulejmenova 1965:76–83) has a cutoff between ST (prothesis) and Sm (variation), and also between Sn (variation) and Sl (anaptyxis). Farsi has its cutoff between Sl and SR (which Hindi as described by Bharati also differentiates).7 For the other languages identified by Fleischhacker, the information is sparser but still consistent with 1.

The scale of splittability in 1 is expected given Fleischhacker’s speculation about perceptual breaks: the SC2 clusters further to the right in the scale have a C2 with stronger formant structure, so the break between S and C2 should be stronger. Or, to take a slightly different view, the more sonorous C2 is, the more vowel-like it is, and thus the more the transition from S to C2 is already similar to a transition from S to a vowel.

The influence of C1 is less clear. Fleischhacker finds only three languages that show a difference between XC2 and YC2, with C2 held constant. One is Farsi, where stop-

---

6 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’s.

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


clusters split (pelastik ‘plastic’, Shademian 2002), but sibilant-l clusters show prothesis. This suggests that TC2 is more splittable than SC2. Similarly, Wolof splits all TL and TR clusters, but shows variation for SI and SR. In Kirgiz, as discussed in detail in Gouskova 2003, there is at least one TC2-SC2 pair, kv (anaptyxis) vs. zv (prothesis), and in general, lower sonority of C1 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 ST but also rt, lb, lv, zv, and mn—all undergo prothesis. The rising-sonority clusters include not just Sm, Sn, Sr, and 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, one would need data from a language with a rich cluster inventory and some cluster-splitting phenomenon that does not create a heterosyllabic C1,C2 sequence, such as C2 deletion.

The Tagalog data to be discussed in this article bear only on SC and TC clusters. We can incorporate the Farsi and Wolof facts into the splittability scale by adding a second dimension, as in 2.

---


* Karimi (1987) and Shademian (2002) both state that prothesis occurs for all SC clusters, but do not investigate SR. Fleischhacker’s data on fr and sw come from Shademian herself.

* A third pattern is vocalization of w, as in [su?et] ‘sweat’ (Karimi 1987:311).

* Bharati states that Sr is usually left intact (i.e. not nativized at all), but that if the cluster does undergo epenthesis, the epenthized form given is Svr.

* Fleischhacker states that the speaker she consulted displayed variation for sibilant-liquid clusters, as in [eslepni] ‘Sleipnir’ but [solovaki] ‘Slovakia’. Only one example is given for Sr, however: [s?i lan?ka] ‘Sri Lanka’. 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, seaeta ‘sweater’). (Data from Breen, n.d.)
(2) $ST < Sm < Sn < Sl < SR, SW$

\[ \bigwedge \bigwedge \]

$TI \quad TR$

Fleischhacker (2001b, 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 Sterlade 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 1954 [1910]:147–48; see also Cairns & Feinstein 1982, Broselow 1992b), copies only the first consonant of most clusters, as in fai-frais ‘tempt-PRETERITE’, gai-grôt ‘weep-PRETERITE’, and sai-slép ‘sleep-PRETERITE’. The clusters st and sk, however, are copied in their entirety, as in ga stai stald ‘possess-PRETERITE’ and skai-skâib ‘sever-PRETERITE’. 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 of English imperfect puns—puns juxtaposing two forms that are not perfect homonyms. Puns like blown apart ~ Bonaparte, 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 $C_1C_2V\ldots \sim C_1V\ldots$, 40 percent 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 $C_1C_2V\ldots \sim C_1V\ldots$, only 26 percent 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, for example, into nasals, liquids, and glides.) Looking at puns of the form $C_1C_2V\ldots \sim C_2V\ldots$, by contrast, such as Stabitha ~ Tabitha, Fleischhacker finds that they occur as often as expected for TL, SL, and ST. The lack of distinctions found among $C_1C_2V\ldots \sim C_2V\ldots$ 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. Fleischhacker also gives evidence for parts of 2 from poetic alliteration, following Kuryłowicz 1971 and Broselow 1992b. In early Germanic (for which Fleischhacker

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

10 The 1,964 puns in Fleischhacker’s corpus come from a book of puns (Crosbie 1977), two books of product slogans (Sharp 1984, Urdang & Robbins 1984), and assorted media sources.

11 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 of juxtaposing the explicit blown-aparte (the name that usually follows Napoleon).

12 For puns of the form $C_1C_2V\ldots \sim C_1VC_2V\ldots$, ..., 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 $C_1C_2V\ldots \sim VC_1C_2V\ldots$ (steamed ~ esteemed) occurs.
cites Kuryłowicz), ST clusters alliterate only cohesively—that is, with themselves; for example, a word beginning in st, such as Old English stan ‘stone’, can alliterate only with words beginning in st, not with words beginning in sV or, say, sp. Words beginning in other C1C2 clusters, however, alliterate with any word beginning in C1: brim ‘sea’ allitrates with beorgas ‘hills’ and blitcan ‘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).13

Splitting in epenthesis, reduplication, the C1C2V... ~ C1V... puns above, alliteration, and VC infixation has the property that if C1C2 is split, C1 becomes vowel-adjacent (C1V...), as summarized in 3. Fleischhacker proposes that in all of these cases there is a preference to keep the two related forms (foreign word and loan, base and reduplicant, etc.) perceptually similar.14

<table>
<thead>
<tr>
<th>(3)</th>
<th>UNSPLIT</th>
<th>SPLIT</th>
</tr>
</thead>
<tbody>
<tr>
<td>epenthesis</td>
<td>C1C2V... (foreign word)</td>
<td>C1VC2V... (adapted)</td>
</tr>
</tbody>
</table>
| reduplication | C1C2V... (base) | C1V... (repet)
| pun | C1C2V... (one member of pun pair) | C1V... (other member) |
| alliteration | C1C2V... (one member of allit. pair) | C1V... (other member) |

Assuming the scale of perceptual distance (Δ) shown in 4, splitting should be most likely when the difference Δ(C1C2,C1V) is small, as in 5, which restates one dimension of 2.

(4) Δ(C1T, C1V) > Δ(C1m, C1V) > Δ(C1n, C1V) > Δ(C1l, C1V) > Δ(C1r, C1V),
Δ(C1W, C1V)

(5) LEAST SPLITTABLE CT Cm Cn Cl CR CW 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 klla ~ ka 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

13 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 one can straightforwardly determine whether a given cluster allitrates 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.

14 In the case of loan adaptation, the preference plays out only in those speakers who have access to the foreign source form.
syllable pairs that delete C₂, subjects discriminated T-sonorant pairs (e.g. kla ~ ka) and S-sonorant pairs (e.g. sla ~ sa) more slowly than ST (ska ~ sa), supporting the view that {S/T}LV and {S/T}V are more similar than STV and SV. 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[a]luck), 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[p]ar ~ s[a][pʰ]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 (pluck ~ [a]pluck), by contrast, ratings seem to be flat across the cluster types.

To summarize Fleischhacker’s findings, a group of phonological and parphonological phenomena—epenthesis, reduplication, punning, and alliteration—display a cross-linguistic 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 cross-linguistic pattern. As in the nasal-assimilation example in §1, one possible explanation is that the phonetics are inside the mind of the speaker: speakers are able to determine how similar a CᵢC₂-C₁V pair is, and are biased to keep pairs such as foreign word and loan, base and reduplicant, and so forth, similar. This would follow Steriade’s (2001a, b) 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 C₁C₂-initial foreign word as having a vowel between the two Cs if C₂ 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 reduplicates, 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

---

15 See Fleischhacker for discussion of CᵢC₂a ~ C₁aC₂a pairs, with schwa epenthesis.
16 Fleischhacker (2001a) and Andersen (1972:36) point out that modifications of this type are common in casual, emphatic English (e.g. puh-leeze ‘please’, kee-rist ‘Christ’).
17 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.
is no convention that makes surgeon ~ sturgeon an illegal pun—but 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 et al. 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-al-ab or s-al-tab would be possible for stub, and pl-ak-ænæot or p-ak-lænæot 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 inflexion 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.

<table>
<thead>
<tr>
<th>REDUPLICATION (Fleischhacker)</th>
<th>ST</th>
<th>&gt;</th>
<th>T-liquid</th>
</tr>
</thead>
<tbody>
<tr>
<td>PUNS (Fleischhacker)</td>
<td>ST</td>
<td>&gt;</td>
<td>S-liquid &gt;</td>
</tr>
<tr>
<td>ALLITERATION (Fleischhacker, Minkova)</td>
<td>ST &gt;</td>
<td>SN &gt;</td>
<td>SI &gt;</td>
</tr>
<tr>
<td>DISCRIMINATION EXPERIMENT (Fleischhacker)</td>
<td>ST,</td>
<td>S-sonorant</td>
<td>&gt;</td>
</tr>
<tr>
<td>SIMILARITY-RATING EXPERIMENT (Fleischhacker)</td>
<td>ST &gt;</td>
<td>SN &gt;</td>
<td>S-liquid, SW,</td>
</tr>
<tr>
<td>INFIXATION GAME (Pierrehumbert &amp; Nair)</td>
<td>ST</td>
<td>&gt;</td>
<td>S-liquid, T-liquid</td>
</tr>
</tbody>
</table>

Table 2. Summary of cluster distinctions. §2.1.

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 & Otanes 1972, French 1988, Prince & Smolensky 2004 [1993], McCarthy & Prince 1993).

(6) bago18 ‘new’ bumago ‘to change’

Native words in Tagalog do not have initial consonant clusters (except for some stop-glide clusters created by optional syncope; see §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).19

(7) ‘graduate’ gumraduate ~ grumaduate
‘protect’ pinrotekta-han ~ prinotekta-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

18 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.

19 There is also a rarer pattern, gumraduate, pinrotekta-han; see §7.1 for discussion of epenthetic vowels, found especially in older loans.
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 inflexion, and whether these follow the crosslinguistic pattern of §2.1. The data to be discussed in §3 come from established loan clusters, and those in §4 come from poorly attested clusters. In both cases, speakers’ treatment of clusters docs 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 World Wide Web. It was constructed as follows. First, a smaller corpus, generously supplied by Rosie Jones (and derived from Ghani et al. 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.

\[(8) \text{STRING} \quad \text{GLOSSES}
\]

kami pangulo ‘we (exclusive)’ ‘president/chief’
lalo parang ‘more/much’ ‘for.LINKER’
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 one thousand queries per day, with each query returning a maximum of ten URLs (Web page addresses); if a query produces more than ten results, only ten are returned at a time and each request for the next ten counts as another query. Thus, a theoretical maximum of ten thousand URLs can be retrieved per day, but the typical number is approximately five thousand, since not all queries return the full ten URLs. Because each Google search returns at most one thousand 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 one thousand.

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 new URL, 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 twenty million words of Tagalog. In a random sample of one hundred pages, twenty-four are blogs, twenty-one are discussion forums, thirteen are newspaper articles, nine are Bible verses, five are press releases, four are nongovernmental organizations’ and social clubs’ sites, and the remaining twenty-four 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.\(^{20}\)

\(^{20}\) Magbabalaod and magbabalaud are probably intrusions from Cebuano. Like the CorpusBuilder software, this method has difficulty keeping out text from other Philippine languages.
This file can then be searched for regular expressions corresponding to potentially infixed forms, such as [ptk]jn[lr][aeiouwy] (p, t, or k followed by n, 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 infixed forms, such as the proper name mckinley.

The initial clusters that have been borrowed into Tagalog upcenthized are almost exclusively C-glide and stop-liquid.21 (As discussed in §4, SC clusters other than s-glide normally undergo prothesis, so that the stem is no longer cluster-initial.) But one prediction made by Fleischhacker’s perceptual account can still be tested. 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-glides was found to be more splittable than sibilant-liquid.

Figures 1 and 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 §7.1 for examples of reduplicated, infixed 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 one 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. In Fig. 1, for example, for the stem practice, there are sixteen tokens total with in, six of prinactice/prinaktis (variant spellings), and ten of pinpractice/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.

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 toward 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

21 There are some loans beginning in nasal-glide or liquid-glide (mw, my, ny, ly), but no infixed examples 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.
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 the two figures, the difference between stop-*r* and stop-*w* is significant by one-tailed Fisher’s exact test ($p < 0.05$ for *in*, $p < 0.002$ for *um*). Differences within the same cluster type across the two figures are not significant.

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, $C_1$ (all stops combined), $C_2$, 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 ($F(3,159) = 10.56, p < 0.0001$) and participates in no significant interactions. The significant pairwise differences ($p < 0.05$ 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), $F(1,161) = 46.84, p < 0.0001$: reduplicated words are less likely to undergo splitting (see
There is a possible etymological confound. 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. Figures 3 and 4 show the results for Spanish-origin 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 figure is stop-r vs. stop-w ($p < 0.05$ for in, $p < 0.005$ for um). There are no significant differences between the two figures. For stop-l and stop-r with each infix, there are no significant differences between Spanish-etymology words and English-etymology words.

As a referee pointed out, this lack of Spanish/English difference suggests that fine details of the source-language form are of little importance. Spanish and English $r$ are very different phonetically ([r] vs. [l]), for example, so one might expect CR clusters from the two languages to be treated differently, if speakers are comparing the source-language form to the infixed form. (Because Spanish-Tagalog contact is now limited,

§7.1 for examples of reduplicated, infixed words). Whether the first syllable of the stem is stressed has a significant main effect, $F(1,161) = 4.82$, $p < 0.01$, 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 (1996), though they describe data that involve infixation with epenthesis. Finally, C1 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.

Thanks to participants in the University of California Berkeley linguistics colloquium for pointing this out.

24 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 transforma 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 translado were excluded from the Spanish-origin counts, but ambiguous cases such as transforma were included.
comparisons to Spanish forms would have had to take place long ago and be fossilized in the word’s contemporary behavior.) Spanish and English l are also quite different following a word-initial voiceless stop. The reader might expect Spanish [pl] to be more splittable because the [l] in p-in-l... is similar to the original; whereas English [pl] 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 infixed form. One might also expect different treatment for Spanish- and English-origin loans if infixation behavior results from variable misperception of cluster-initial foreign words (see §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 n. 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 five 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.

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 n. 19, some older loans from Spanish have an epenthetic vowel, as in palantza ‘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 §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 §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 C1(C2)V1-C1C2V1 only). For stop-liquid clusters, there are two options: simplified (e.g. mag-tatrabaho ‘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-byaha ‘will travel, mag-ke-kwento ‘will narrate’) or, more commonly, by having a reduplicant vowel that corresponds in color to the glide (mag-bi-byaha, mag-ku-kwento). The two simplified options together occur about two-thirds of the time, and the fully copied option (mag-bya-byaha, 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 prosthesis when borrowed into Tagalog. For example, ‘scan’ is normally pronounced [ʔiskən], and the infix is placed before the prothetic vowel ([ʔumiskən], cf. native [ʔumawit] ‘to sing’—see n. 27 for discussion of the glottal stop’s status), so that the issue of whether to split the cluster does not
arise. Some speakers use nonprothetized forms such as [skan], but these very rarely occur with infixation. In the corpus, there were only twenty-four tokens, seventeen of them from a single type, the nickname of a sports team, which may have originated in a speech error.  

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, thus allowing participants to be located anywhere in the world while completing it. I hoped that many of the survey participants would be living in the Philippines, and thirty-five (out of sixty-two 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 fourteen 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 Fig. 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.

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 non-split option is simply the mirror image). Splitting was seldom chosen for s-stop clusters (on the left), but was usually chosen for sw clusters (on the right).

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 Figs. 1, 2, and 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%

25 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 web site, at eskumor.sitesledd.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 monicker of the Bloomfield Basketball team before it was disbanded in 2006.’ The other tokens are scrinutinize, iskinetch (from sketch—this word may have the prefix i or be formed similarly to eskumor), silnice, silnow, sprinayan (from spray, with the suffix -an), spinray-paint, sitwalk and stino-talk, strumay, and struming.

Unprothetized s-initial clusters are somewhat more common with reduplication than with infixation (e.g. pag-sno-snorkel ‘snorkeling’, mag-isi-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.
Figure 6. Sample survey page: forced-choice task and ratings task.

Figure 7. Results of forced-choice task: rate at which split option was chosen, for each cluster.
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- 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-llr and sibilant-llr, except to note that stop-llr clusters are well attested with infixation among existing loans, whereas infixation of sibilant-llr is basically novel—perhaps, for some unknown reason, cluster novelty increases 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).

Figure 8. Results of ratings task: mean ratings for both options, for each cluster. (Error bars indicate 95% confidence interval.)

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, shr, sw), and no between-subjects factors, with rating difference as the dependent variable, F(5, 50)26 = 5.80, p < 0.001, with a (vacuous) Huynh-Feldt correction to degrees of freedom. Paired (by participant) t-tests can then be performed on

The full ANOVA is applied only to subjects who rated all six cluster types. Pairwise comparisons include a few more subjects.
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.

<table>
<thead>
<tr>
<th></th>
<th>RATING DIFFERENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CHOICES</td>
</tr>
<tr>
<td>sm</td>
<td>p = 0.0892</td>
</tr>
<tr>
<td></td>
<td>p = 0.0049</td>
</tr>
<tr>
<td>sn</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>p = 0.0019</td>
</tr>
<tr>
<td></td>
<td>p = 0.0045</td>
</tr>
<tr>
<td>sl</td>
<td>p = 0.0021</td>
</tr>
<tr>
<td></td>
<td>p = 0.0031</td>
</tr>
<tr>
<td></td>
<td>p = 0.7305</td>
</tr>
<tr>
<td></td>
<td>p = 0.8392</td>
</tr>
<tr>
<td>Sr</td>
<td>p = 0.0003</td>
</tr>
<tr>
<td></td>
<td>p = 0.0003</td>
</tr>
<tr>
<td></td>
<td>p = 0.1539</td>
</tr>
<tr>
<td></td>
<td>p = 0.0666</td>
</tr>
<tr>
<td></td>
<td>p = 0.0945</td>
</tr>
<tr>
<td>sw</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>p = 0.0030</td>
</tr>
<tr>
<td></td>
<td>p = 0.0021</td>
</tr>
<tr>
<td></td>
<td>p = 0.0018</td>
</tr>
<tr>
<td></td>
<td>p = 0.0953</td>
</tr>
</tbody>
</table>

Table 3. Significance of pairwise differences between clusters in survey results. (Cells are shaded when p < 0.05 for at least one measure.)

If the survey results are taken 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.

![Figure 9](image_url)

Figure 9. Significant differences in splittability, from survey data.

Although the predicted distinctions among sn, sl, and shr 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,b, 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 \(r\) becomes surface \(y\), and \(\text{FAITH2}\) is violated when underlying \(z\) becomes \(w\), and \(\Delta(x,y) > \Delta(z,w)\), then, by default \(\text{FAITH1} \gg \text{FAITH2}\) (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 Fleischhacker proposed (2001a) given in 4 is repeated as 10, with \(S\) substituted for \(C_1\) (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) \(\text{DEP}-\text{V/S}_T \gg \text{DEP}-\text{V/S}_m \gg \text{DEP}-\text{V/S}_n \gg \text{DEP}-\text{V/S}_1 \gg \text{DEP}-\text{V/S}_R \gg \text{DEP}-\text{V/S}_W\)

\(\text{DEP}\) constraints (McCarthy & Prince 1995) penalize insertion of segments. These are context-sensitive \(\text{DEP-\text{V}}\) constraints, which penalize insertion of a vowel in a particular context, such as between a sibilant and a stop (\(S_T\)) as in /sparta/ \(\rightarrow, [\text{separta}].\)

By ranking \(\text{LEFT-ANCHOR}\) (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 cutoff point for cluster splitting. Additional markedness and faithfulness constraints determine which unsplit clusters are adapted faithfully and which receive a preceding epenthetic vowel. Prince and Smolensky’s (2004 [1993]) *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 cutoff constraint is not \(\text{LEFT-ANCHOR}\) but rather a markedness constraint against consonant clusters. The tableaux in 12 illustrate the analysis for a language that prothesizes sibilant-stop clusters, and epenthesis sibilant-\(l\) clusters.

(12) Schematic analysis of asymmetric epenthesis pattern

<table>
<thead>
<tr>
<th>source word [spV...]</th>
<th>*COMPLEX</th>
<th>(\text{DEP-\text{V/S}}_T)</th>
<th>LEFT-ANCHOR</th>
<th>(\text{DEP-\text{V/S}}_1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. spV…</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. sipV…</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. spispV…</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>source word [slV...]</th>
<th>*COMPLEX</th>
<th>(\text{DEP-\text{V/S}}_T)</th>
<th>LEFT-ANCHOR</th>
<th>(\text{DEP-\text{V/S}}_1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>d. slV…</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. spisV…</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>f. islV…</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

In order to extend this account to similar patterns in reduplication, imperfect puns, and alliteration, Fleischhacker (2001b) introduces an additional family of default-ranked contextual \(\text{MAX}\) constraints, which penalize deletion of segments (McCarthy & Prince 1995), shown in 13 (cf. Fleischhacker’s \(\text{DEP}\) 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 (gai-grot). 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).

\(13\) \(\text{Max} \ T/S \ V \gg \text{Max} \ m/S \ V \gg \text{Max} \ n/S \ V \gg \text{Max} \ l/S \ V \gg \text{Max} \ r/S \ V \gg \text{Max} \ W/S \ V\)

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 \gg \text{Contig-Sii} \gg \text{Contig-Si} \gg \text{Contig-SI} \gg \text{Contig-SR} \gg \text{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-SI). For the contextually sensitive Contig family to reflect the similarity claims in 10, the context in which the Contiguity constraint applies must be further specified, 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/. . . \gg \text{Contig-Sm/V} . . . \gg \text{Contig-Sn/V} . . . \gg \text{Contig-Si/V} . . . \gg \text{Contig-SR/V} . . . \gg \text{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, \(\text{Map-S}_{1}S_{2}^{A}(X,B,C,Y,D)\): an \(X\) in the environment \(A,B\) in string \(S_{1}\) must not correspond to a \(Y\) in the environment \(C,D\) in string \(S_{2}\). For example, \(\text{Map}(S^{T}, S^{Y})\) forbids a sibilant that is followed by a stop from corresponding to a sibilant that is followed by a vowel. The hierarchy \(\text{Map}(S^{T}, S^{Y}) \gg \text{Map}(S^{m}, S^{Y}) \gg \text{Map}(S^{o}, S^{Y}) \gg \text{Map}(S^{l}, S^{Y}) \gg \text{Map}(S^{s}, S^{Y}) \gg \text{Map}(S^{w}, S^{Y})\) would cover the three hierarchies in 11, 13, and 15. Donca Steriade suggested to me a less radical move that would also cover all three cases: C-Contiguity(ST), 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 a single family of constraints that covers all the cases (epenthesis, infixation, partial reduplication, etc.) can be formulated.

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 \(\text{Complex}^{*}\), 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 2004 [1993]), 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.\(^{27}\)

---

\(^{27}\) Ross (1996) attempts to repair the NoCoda analysis by adding variably ranked \(\text{Complex}^{*}\), which would prefer g-um-rad.wet. If, however, \(\text{Complex}^{*}\) 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.
(16) Schematic analysis of asymmetric infixation pattern

<table>
<thead>
<tr>
<th></th>
<th>ANCHOR-STEM</th>
<th>CONTIG-ST/V...</th>
<th>LEFTMOST</th>
<th>CONTIG-SI/V...</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. in + spin</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. sinpin</td>
<td></td>
<td></td>
<td>s</td>
<td></td>
</tr>
<tr>
<td>c..pair spinin</td>
<td></td>
<td></td>
<td></td>
<td>sp</td>
</tr>
<tr>
<td>um + spin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. umslip</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. pair sumlip</td>
<td></td>
<td></td>
<td>s</td>
<td></td>
</tr>
<tr>
<td>f. slumip</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

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 et al. 2003), derive idealized outputs shown in Figure 10 (cf. the survey results in Fig. 7).

6. Alternatives. I 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 C1C2 and C1V varies depending on C2, and that they apply this knowledge so as to maximize the similarity of infixed and unfixed words.

As mentioned in §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 experi-

about which clusters should split more often. See the discussion of cluster markedness in §8.2. Moreover, language-internal evidence requires that *COMPLEX >> NOCODA, since word-internal clusters are syllabified heterosyllabically (ak.lat ‘book’).

The reader might object 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 infixed 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.

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-Ca/u, we expect labusaw ‘made turbid’, to be infixed 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).

28 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.
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 more than 20% of the time, as shown in Fig. 10) while Contig-S∩/V is only somewhat likely to outrank Leftmost (producing splitting of sm almost half the time), and Contig-S∩/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.](image)

For our purposes, the set of possible learning theories can be partitioned 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

---

29 Re ‘ordinary methods’: as Idaardi (2007) points out, there is no such thing as a purely data-driven learner—we cannot 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.
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 isoC 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 infixed words of the form s-in-C. . . /isoC-in-. . . or s-um-C. . . /isoC-um-. . . to establish the rate at which the infix splits each type of isoC 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): §§7 and 8.

d. Give evidence that learners do not in fact have access to (b): the extremely low corpus frequency of infixed isoC-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 of 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 §7, and accounts that endow learners with language-specific knowledge other than a similarity bias (§8).

Although my proposal 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 as ‘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.30

30 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.
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].[31] 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 $C_1C_2$ cluster.[32] To explain cluster differences, we could plausibly assume that greater sonority of $C_2$ 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 §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 inflexion with reduplication.

In native words, when inflexion 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-hago ‘is changing’, from the stem hago ‘new’. When this construction is applied to a cluster-initial loan, several variants are possible. Examples are shown in Table 5, with corpus frequencies.

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-uwa-guwapo], [p-in-o-poro-

<table>
<thead>
<tr>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ONSET COPIED, NOT SPLIT BY INFIX</strong></td>
<td><strong>ONSET COPIED, SPLIT BY INFIX</strong></td>
<td><strong>ONSET SIMPLIFIED, $C_2$ SKIPPED</strong></td>
<td><strong>ONSET SIMPLIFIED, $C_2$ VOCALIZED (if $C_2$ is glide)</strong></td>
</tr>
<tr>
<td>0</td>
<td>g-um-wa-gwapo</td>
<td>0</td>
<td>g-um-a-gwapo</td>
</tr>
<tr>
<td>8</td>
<td>s-um-wa-sweldo</td>
<td>7</td>
<td>s-um-u-sweldo</td>
</tr>
<tr>
<td>kw-in-e-kwenta</td>
<td>0</td>
<td>k-in-e-kwenta</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>b um ya byahe</td>
<td>0</td>
<td>k-in-e-kwenta</td>
</tr>
<tr>
<td>pr-in-o-problema, 28</td>
<td>0</td>
<td>b um a byahe</td>
<td>0</td>
</tr>
<tr>
<td>pr-in-u-problema</td>
<td>249</td>
<td>b um i byahe</td>
<td>4</td>
</tr>
<tr>
<td>pr-in-o-promote, 11</td>
<td>1</td>
<td>pr-in-u-promote</td>
<td>6</td>
</tr>
<tr>
<td>pr-in-u-promote</td>
<td>0</td>
<td>p-in-re-prepare</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>p-in-ri-pritu</td>
<td>32</td>
<td>NA</td>
</tr>
</tbody>
</table>

**Table 5.** Corpus-attested variants for reduplication + inflexion, with token frequencies.

---

31 Cena (1979) assumes that splitting of a loan cluster by the inflex (and partial reduplication) results from an extra vowel, but in the examples he considers the vowel is robust (and spelled).

32 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 the extra vowel is sometimes 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 reduplicate makes sense only if the stem is treated as *bya.he*, not *bi.ya.he.*
blema], and so forth, 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 svarabhatki 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. 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.

7.3. ‘Insert infix before first x’. Finally, a referee 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’), predomminating over infixation for l and y, and rare for h and w. When the prefix [i] 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): [i]-ni-?abot ‘handed to’, cf. [i]-in-?abot ‘reached for’; in is still used, however, for other consonants ([?i-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

33 Thanks to Colin Wilson and Christian Uffmann for raising this possibility.

34 At a referee’s suggestion, this idea was also implemented using not observed frequencies of the clusters, but rather 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.
\(X\)', where \(X\) is a feature matrix.\(^{35}\) The simplest and most reliable generalization has \(X = [+\text{syllabic}]\) ('insert the infix before the first vowel'), but others have some support too. \(X = [-\text{consonantal}]\), for example, does fairly well, because it is true of words like \(b-um-ili\) ([i] is [-cons]), and also of words like \(n-i-yakap\) (\(y = [j]\) is [-cons]), and false only of words like \(y-um-akap\) (the rarity of words like \(w-um-agayway\) helps boost this generalization). \(X = [+\text{sonorant}]\) does less well, but still not too badly. It is true of words like \(b-um-ili\) and \(n-i-lagay\), but false of words like \(l-um-akad\). One of the worst generalizations, with \(X = [-\text{sonorant}]\), is never true (assuming [\(h\)] and [\(\theta\)] are [+sonorant]): \(n-i\) never occurs with obstruent-initial stems. If speakers then apply these generalizations in the survey task, \(s-\text{in-werte}\), which obeys 'insert the infix before the first [-cons]', should be rated higher than \(s-\text{in-lip}\), which obeys the weaker 'insert the infix before the first [+son]', which should in turn be rated higher than \(s-\text{in-top}\), which obeys the always-false generalization 'insert the infix before the first [-son]'.

There are many values for \(X\), 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 is the \textsc{maximum entropy} framework. A full explication of this framework is beyond the scope of this article; see Goldwater & Johnson 2003 on applying maximum entropy to \textsc{OT}-like constraint weighting. In the course of learning, each generalization is assigned a weight (or in this case, the inverse of each generalization

\(^{35}\) Generalizations of the type 'insert the infix \textsc{after} the first \(X\)' have little hope of working, because most of the generalizations would fare so badly on the native data: \(X = [-\text{syllabic}]\) does well, but \(X = [-\text{continuant}]\), for example, which must receive a large weight in order to favor \textsc{st-in-op} over \textsc{s-in-top}, is falsified by abundantly many forms such as \(l-um-akad\), not \(\*lak-um-ad\).
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.36 The training data—the frequency of ni, in, and um for stems beginning in CV. . . , for all values of C and V—were derived from a mixture of corpus counts and extrapolations from a database of disyllabic native roots, from English 1986.37

As shown in Figure 12, the resulting set of weighted generalizations manages to distinguish sw from the other sC clusters, splitting it about 5 to 50 percent 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 percent of the time.

![Figure 12. Splitting rates predicted by MaxEnt model.](image)

I conclude that at present none of the data-driven accounts of the survey data works well enough to be accepted. 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 §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

36 I am 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.

37 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. Stems beginning with n and taking the in/ni suffix are difficult to classify: is ninakaw 'be robbed' n-in-akaw or ni-nakaw? There were only seventeen such words in the corpus (nasal-initial roots are underrepresented in Tagalog), and each was counted as half infixed and half prefixed.
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 [kalam], a variant of [kalm] ‘calm’. 38

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.

Loose coordination of a CC cluster could plausibly lead to greater splittability, even in a language that does not have (noticeable) excrescent vowels. Suppose that obstruent-obstruent clusters such as ST are subject to a constraint requiring the release of S to coincide with the target of T (i.e. there is no gap between the two consonants). 39 If that constraint is defined to apply to UNDERLYINGLY adjacent S and T, then it would be violated if an infix splits the cluster. Obstruent-sonorant clusters (i.e. all the other Tagalog clusters examined here) would not be subject to this constraint, and so lesser splittability of ST as compared to all the other clusters is predicted. 40

Looking at differences within the sonorants, Hall finds that in most languages not all sonorants trigger a svarabhakti vowel, and she proposes the implicational hierarchy in 18.

\[
\begin{align*}
&\text{(18) obstruents} < \text{glides, nasals (within which } m < n) < r < l < r, \gamma < \text{gutturals} \\
&\text{LEAST LIKELY TO TRIGGER} \quad \text{SVARABHAKTI} \quad \text{MOST LIKELY TO TRIGGER} \quad \text{SVARABHAKTI}
\end{align*}
\]

This is similar to Fleischhacker’s hierarchy for epenthesis in SC clusters, which raises the possibility that the hierarchies really both follow from the same cause, whether articulatory or perceptual.

38 Warner and colleagues (2001) argue that the schwa in forms like [kalum] results from a separate vocalic gesture, because the articulation of [l] in forms like [kalum] patterns more with [l] before underlying schwa than with [l] in forms like [kalum]. Hall counters that the articulatory difference between the [l] articulations in [kalum] and [kalum] could result from the timing difference, rather than from a true epenthesis.

39 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 C1 coincides with the target of C2, there is no interval of open vocal tract between the two consonants.

40 This is not exactly faithful to Hall’s account of svarabhakti vowels. She proposes a general constraint, applying to all consonants, requiring alignment of C1’s release to C2’s target, and a specific constraint for obstruent-sonorant clusters requiring obstruent C1’s center to be aligned with sonorant C2’s onset, a configuration that results in an excrescent vowel. These two constraints would both be violated by infixation into an obstruent-sonorant cluster.
There is one definite mismatch between Hall’s hierarchy for svarabhakti and Fleischhacker’s for epenthesis: the place of glides within the hierarchy. In this respect, the 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 Si clusters are not, the rhotic is a tap, [r] (Isfari lanjal ‘Sri Lanka’, Shabnam Shadamian, 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 Si 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 [l], 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, one 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; one 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. Over all, 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 infixed word. 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. But this seems 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.41

41 As mentioned in n. 27, 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.
(20) **least often split**

sibilant-stop (ST) sib.-nasal (Sm, Sn) sib.-liquid (Sl, Sr) stop-liquid (Tl, Tr) stop-glide (TW)

**most often split**

sib.-glide (SW)

**most marked**

There are a few criteria that could be used to determine which clusters are more marked. Crosslinguistically, it has been claimed that the greater the sonority increase from C\textsubscript{1} to C\textsubscript{2}, the less marked is the onset cluster C\textsubscript{1}C\textsubscript{2} (e.g. Greenberg 1978, Selkirk 1984).\textsuperscript{42} This would mean that TW is less marked than T-liquid and that the SC clusters toward the right in 20 are less marked than those toward the left. Steriade’s (1999) theory of consonant cuing claims that consonant clusters are marked because of C\textsubscript{1}’s reduced perceptibility: C\textsubscript{1} lacks a following vowel or sonorant whose formants it can alter, and lacks a release burst. This predicts that greater sonority of C\textsubscript{2} should reduce markedness: again, TW should be less marked than T-liquid, and the SC clusters toward the right should be less marked than those toward 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. Consider first 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 §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\textsubscript{1}VC\textsubscript{2} and C\textsubscript{1}C\textsubscript{2} when C\textsubscript{2} is a glide (and V matches it in color, that is, backness and rounding), but not when C\textsubscript{2} is a liquid, no matter what the intervening vowel.\textsuperscript{43}

\begin{align*}
(21) & \text{[pi}j\acute{a}k] \sim \text{[pi}j\acute{a}k] ‘\text{squawk’} \\
& \text{[bu}\text{w\acute{a}n}] \sim \text{[bw\acute{a}n]} ‘\text{moon’} \\
& \text{[pr}\acute{\text{o}}k] \sim \text{*pr}\acute{\text{o}}k] ‘\text{district’}
\end{align*}

This suggests that TW is less marked than T-liquid, though it is also possible that similarity preservation is at work here. Since TVW is highly similar to TW—especially if V matches the glide in color—deletion of 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\textsubscript{1},C\textsubscript{2}, avoiding a complex onset. Evidence for this syllabification comes from speakers’ intuitions (Schachter & Otanes 1972) and from stress facts. Stress (sometimes characterized as length—see Schachter & Otanes 1972, 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).

\begin{align*}
\text{(22) (a)} & \quad \text{unsuffixed forms} & \quad \text{(b) suffixed forms} \\
\text{open penult: penultimate or final stress} & \quad [\text{bi}r\acute{\text{u}}?] ‘\text{joke’} & \quad [\text{bi}r\acute{\text{u}}?-\text{in}] ‘\text{to joke’} \\
& \quad [\text{ta}n\acute{\text{o}}p] ‘\text{question’} & \quad [\text{ta}n\acute{\text{n}}-\text{in}] ‘\text{to question’} \\
\text{closed penult: final stress only} & \quad [?i\acute{\text{k}}.li?] ‘\text{shortness’} & \quad [?i\acute{\text{k}}.li.?-\text{an}] ‘\text{to shorten’} \\
& & \quad \text{(English 1986)}
\end{align*}

\textsuperscript{47} Steriade (2004), however, proposes that in Latin, CW clusters are more marked than other clusters.

\textsuperscript{43} 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).
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 in 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 clusters. Again, this makes the wrong prediction for the splitting facts.

(23) a. [rén.da] ‘rein’ [ren.da.h-án] ‘to rein’ (Spanish rienda)

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-radunate. If this force is responsible for differences in cluster splittablility, then we expect that C1,C2 should be more splittable the less marked a nasal-C2 cluster is. Again, this is the opposite of what happens.

(24) LEAST OFTEN CREATED
    nasal-stop      nasal-nasal      nasal-liquid      nasal-glide
    MOST OFTEN CREATED
    LEAST MARKED
    MOST MARKED

In order to establish nasal-C cluster markedness, one 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 the syllable contact law is interpreted gradiently, so that flat sonority is also worse than rising sonority, and the greater the sonority rise, the worse, then the clusters in 24 become more marked toward the right.

Tagalog-externally, we can compare type frequencies of root-internal nasal-C clusters, shown in Figure 13. 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 (‘kr, 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 C2 = r, because [r] in native words does not occur after a nasal.)

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 and 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

---

44 Counts are from disyllabic native roots found in English 1986.
Figure 13. Type frequencies in dictionary of root-internal CC clusters.

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 in understanding crosslinguistic trends. Berko (1958), for example, showed that English speakers can generalize from existing plural nouns to form new plurals by adding [iz] after sibilants, [s] after voiceless nonsibilants, 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[iz], feet, etc.). But Berko’s results do not tell us whether learners privilege an assimilatory pattern like that in English 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. Crosslinguistic trends are relevant to developing an explanatorily adequate theory only if they tell us something about learner preferences, and as discussed in §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 §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 (2001a,b): 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 C₁-C₂ transition is to a C₁-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 (§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 one’s 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.

These instructions are 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 Fig. 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 + swere, um + sweldo}
FILLER ITEMS
• in + byabe
• um + byabe
• in + bwisit
• um + bwelo
• 1 of {in + flash, in + frame}
• 3 of {in + syuting, in + pweasto, in + block, in + break, um + drive, in + drive, in + dwrewing, um + gabe, um + gwapo, in + create, in + kwento, in + plano, in + promote, in + pwersa, um + pwersa, in + twahabo, um + twahabo}

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 two such errors, or if the participant completed fewer than five items, all data from that participant were excluded.

REFERENCES


Hayes, Bruce; Robert Kirchner; and Donca Steriade (eds.) 2004. *Phonetically based phonology*. Cambridge: Cambridge University Press.


THE ROLE OF PHONETIC KNOWLEDGE IN PHONOLOGICAL PATTERNING


Zhang, Jie; Yuwen Lai; and Craig Turnbull-Sailor. 2007. Revisiting the psychological reality of Taiwanese tone sandhi. Lawrence: University of Kansas, ms.

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


Department of Linguistics
University of California, Los Angeles
3125 Campbell Hall, Box 951543
405 Hilgard Ave.
Los Angeles, CA 90095-1543
[kie@ucla.edu]

[Received 4 April 2005; revision invited 2 October 2005; revision received 22 May 2006; accepted 20 November 2006]