Subcategorization vs. Output Optimization in Syllable-Counting Allomorphy*

Mary Paster
University of California, Berkeley

1. Introduction: ‘Syllable-counting allomorphy’

Phonologically conditioned suppletive allomorphy (PCSA) was only recently overtly acknowledged in the literature (Carstairs 1988) and has not been studied systematically since Mester (1994) first proposed using OT to model it. Syllable-counting allomorphy (SCA) is a particularly common and interesting subtype of PCSA. It was studied and discussed by Kager (1996) but without the benefit of a large database. In this paper, I focus on SCA because it is the subject of Kager’s (1996) claim that PCSA is output optimization.

An example of SCA is seen in Dyirbal (Pama-Nyungan, Australia; Dixon 1972: 42, 224). Among vowel-final stems, disyllabic stems mark ergative with -ŋgu, while longer stems take -gu, as in (1).

(1) yaraŋgu ‘man’ yamaŋniŋu ‘rainbow’
yuguŋgu ‘stick’ duŋnaŋunuŋu ‘from leaves in water’

In this talk, I will introduce Kager’s (1996) Output Optimization approach (§2), summarize data from a cross-linguistic survey of PCSA/SCA (§3), introduce the Subcategorization approach (§4), and conclude by comparing the two approaches (§5).

2. The Output Optimization approach

Kager (1996: 170) claims that ‘Syllable-counting allomorphy is an output-oriented phenomenon,’ resulting from the Emergence of the Unmarked (TETU; McCarthy & Prince 1994). In this approach, the morphology supplies multiple suppletive allomorphs of an affix in candidates in an OT tableau, and phonological foot structure and parsing constraints select the appropriate output. This approach falls under McCarthy & Prince’s (1993a,b) ‘P >> M’ ranking schema, where phonological effects in morphology are modeled by ranking phonological (P) over morphological (M) constraints. In PCSA, the relevant M constraint is one enforcing uniform marking of morphological categories. One prediction of this approach is that syllable-counting allomorphy should serve to optimize words with respect to phonological constraints that are independently motivated elsewhere in Universal Grammar.

Kager’s argument is based on Estonian (Mürk 1991), where the Genitive plural, Partitive singular, and Partitive plural exhibit SCA, as seen below (examples are from Kager 1996: 157, 168).

(2) a. Genitive plural: -te occurs with even-syllable stems; -tte occurs with odd-syllable stems
   visa-te (no gloss) para'ja-te (no gloss)
   atmi'rali-te ‘admiral’ raamatu-te ‘book’

b. Partitive plural: -sit occurs with even-syllable stems; -it occurs with odd-syllable stems
   visa-sit (no gloss) para'ja-it (no gloss)
   atmi'rali-sit ‘admiral’ raamatu-it ‘book’

*Thanks to Sharon Inkelas, Larry Hyman, Keith Johnson, David Mortensen, other participants in the UC Berkeley phonetics/phonology group, and the audience at WCCFL 24 for helpful comments.

© 2005 Mary Paster
Cascadilla Proceedings Project
Completed April 9, 2005
c. Partitive singular: -t or -Ø occurs with even-syllable stems; -tt occurs with odd-syllable stems
   'ema-Ø  'mother'    'raamatu-tt  'book'
   'tele'oni-Ø  'telephone'  'numis'maatiku-tt  'numismatics'

Kager claims that these apparent syllable counting effects in Estonian result from foot parsing
considerations rather than from true 'counting'. Kager proposes (1996: 162-163) the following
constraints for Estonian:

(3)  FT-BIN: Feet are binary under syllabic or moraic analysis.
PARSE-2: One of two adjacent stress units (μ, σ) must be parsed by a foot.
ALIGN-HD-L: Align (PrWd, L, Head(PrWd), L)
ONSET: Syllables must have onsets.
ALIGN-ST-R: Align (Stem, R, Foot, R) (Each right stem edge coincides with a right foot edge)
PK-PROM: Peak(x) > Peak(y) if |x| > |y|.

The correct genitive plural allomorphs for even-syllable (4a) and odd-syllable (4b) stems are selected
via the following constraint rankings (Kager 1996: 164-167).

(4)  Genitive plural

<table>
<thead>
<tr>
<th></th>
<th>/visa, {-te, -tte}/</th>
<th>FT-BIN</th>
<th>PARSE-2</th>
<th>ALIGN-HD-L</th>
<th>ONSET</th>
<th>ALIGN-ST-R</th>
<th>PK-PROM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[vi.(sá-t.te)]</td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>[vi.(sá-t. te)]</td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>[vi.(sá-t.t.e)]</td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Similarly, the partitive plural allomorphs for even-syllable (5a) and odd-syllable (5b) stems are
selected as shown below (Kager 1996: 164-167).

(5)  Partitive plural

<table>
<thead>
<tr>
<th></th>
<th>/visa, {-sit, -it}/</th>
<th>FT-BIN</th>
<th>PARSE-2</th>
<th>ALIGN-HD-L</th>
<th>ONSET</th>
<th>ALIGN-ST-R</th>
<th>PK-PROM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[vi.(sá)-sit]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>[vi.(sá)-sit]</td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>[vi.(sá)-sit]</td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>[vi.(sá)-sit]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
This type of analysis works for Estonian, but we are left with the question of whether SCA is always reducible to general markedness and faithfulness, as Kager claims. In the following section, I address this question via a cross-linguistic survey of SCA.

3. Survey results

In this survey, 22 cases of SCA were found in 16 different languages (listed in Appendix). Approximately 600 languages were surveyed, as part of a larger study of PCSA (see Paster in progress). In 10 of the 22 cases of SCA, the distribution of allomorphs is not predictable from their shape and must be stipulated; this in itself is an important result since it contradicts Kager’s claim that SCA is optimizing. In the remaining 12 cases of SCA, the distribution of allomorphs does relate to their shape. However, in some of these cases, the phonological constraints that would be needed to model the distribution are very specific constraints that cannot be assumed to be part of UG, and therefore these cases still cannot be analyzed as TETU effects.

For example, in Kimatuumbi (Bantu, Tanzania; Odden 1996), monosyllabic verbs mark perfective with -i¶te, while polysyllabic stems take an -i¶- infix (Odden 1996: 51–53), as shown below.

(6) chóla 'draw (inf.)'   áandjika 'write (inf.)'
nj-chól-i¶te 'I have drawn' nj-áandjikke 'I have written'
ti 'chop (inf.)'   bëleka 'bear (inf.)'
nj-tin-i¶te 'I have chopped' nj-bëlijke 'I have borne'
káata 'cut (inf.)'   chëliya 'be late (inf.)'
nj-káat-i¶te 'I have cut' nj-chëliyë 'I have been late'

The distribution could be construed as responding to a requirement that perfective stems have three syllables, though this is not surface-true due to some systematic exceptions. But note that if this were captured using a phonological constraint, it would be language-specific, not part of UG. Therefore, Kimatuumbi perfective allomorphy is not a TETU effect.

3.1 Examples of non-optimizing SCA

As mentioned above, this survey revealed 10 examples of non-optimizing SCA. An example is found in Tzeltal (Mayan, Mexico; Walsh Dickey 1999), where the perfective is marked by -oh with monosyllabic stems, and with -eh elsewhere (examples are from Walsh Dickey 1999: 328–329).

(7)  s-ku¶ë-oh 'she has carried it'  s-ku¶ë-la¶-eh 'she was carrying it repeatedly'
s-ma'ë-oh 'he has hit something'  s-makli¶-ëh 'he has listened to something'
j-i¶-oh 'he has seen something'  s-makli¶-ëh 'he has fed someone'
s-nu'ts-oh 'he has chased something'  h-pak'-anta'ëh 'I have patched it'
j-a¶l-oh 'he has told something'  s-tiku'n-ëh 'he has sent something'
s-jo¶m-oh 'he has gathered it'  s-majli¶-ëh 'he has waited for someone'

[o] and [ë] do not alternate elsewhere (Kaufman 1971: 28), so the allomorphy is probably truly suppletive. Stress in Tzeltal is word-final (Walsh Dickey 1999: 327), so the allomorphy is not stress-conditioned. A constraint banning [ë] in the second syllable has not been proposed for UG, so this is not a TETU effect. Such a constraint in Tzeltal would be stipulative and used only for this phenomenon, so this appears to be a case where we would not want to describe the distribution of allomorphs as phonologically optimizing in any way.

Another such case is seen in Kaititj (Pama-Nyungan, Australia; Koch 1980). In this language, the -η form of the ergative/instrumental/locative suffix occurs with disyllabic stems, while -l occurs with
larger stems (examples are from Koch 1980: 264-266). Note that in these examples, /l/ → [l] when the preceding consonant is apical (Koch 1980: 274), and N is a prestopped apical nasal.

(8) a’ki-ŋ ‘head’ a’lik-i-1 ‘dog’
il’ti-ŋ ‘hand’ a’tuy-i-1 ‘man’
a’N’mi-ŋ ‘red ochre’ a’yirki-1 ‘sun’
ayn’pi-ŋ ‘pouch’ [uNpiri-1] ‘forehead’

There does not seem to be anything ‘better’ about /ŋ/ rather than /l/ as a stressed syllable or second syllable coda, so this seems once again to be a case of non-optimizing SCA.

A third example is found in Zuni (isolate, New Mexico; Newman 1965). In Zuni, singular nouns are marked with -leʔ for monosyllabic stems, and with -nne for polysyllabic stems (examples are from Newman 1965: 24, 56).

(9) ’i-leʔ ‘sinew’ ’homa-nne ‘juniper leaf’
ši-leʔ ‘piece of meat’ ’tena-nne ‘song’

Again, there is nothing to suggest that the distribution of allomorphs is phonologically optimizing here.

Finally, a fourth example of non-optimizing SCA is seen in Dyirbal. As discussed in §1, in vowel-final stems in Dyirbal, disyllabic stems mark ergative case with -ŋgu, while longer stems take -gu (examples are from Dixon 1972: 42, 224 and reproduced from (1) above).

(10) ’yara-ŋgu ‘man’ yama-ni-ɡu ‘rainbow’
yugu-ŋgu ‘stick’ ɖuŋa-ŋunu-ɡu ‘from leaves in water’

McCarthy & Prince (1990) cite this as an example of ‘complementarity’ in allomorphy. However, as McCarthy & Prince acknowledge, consonants are not moraic in this language, so -ŋgu is not really larger than -gu. Furthermore, in comparing two- vs. three-syllable stems, the distribution of allomorphs is exactly wrong as a TETU effect. Stress is initial and alternating in Dyirbal (Dixon 1972: 274-276; final syllables are unstressed). Thus, the ɡ of -ŋgu has the effect of closing unstressed syllables; a TETU effect involving the Weight-to-Stress Principle would have produced the exact opposite distribution of allomorphs.

3.2 Examples of SCA possibly resulting from TETU

The survey did reveal some cases of SCA aside from Estonian that may be analyzed as TETU effects. For example, in Shipibo (Panoan, Peru; Elías-Ulloa 2004), the repetitive is marked by -ribi with even-syllabled stems, and by -riba elsewhere (examples from Elías-Ulloa).

(11) ’yono-ribi-ki ‘commanded again’ yo’mi-tso-riba-ki ‘stole again’
command-again-past steal-again-past
’ka-ma-ribi-ki ‘made him/her go again’ ’ka-yama-riba-ki ‘do not go again’
go-caus-again-past go-neg-again-past

Elías-Ulloa (2004) claims that this allomorphy is driven by sonority, so that /a/ is a good foot head, while /i/ is not (foot-initial syllables are assumed to be heads). The main stress is on the second syllable if closed; otherwise, the first syllable. There is no secondary stress. Elías-Ulloa uses the constraints *i/Head and *a/NonHead to select the correct allomorphs in a TETU analysis.

Note, however, that we do not need to assume that the allomorphy is suppletive. We could say simply that the repetitive suffix is -ribiV/ with an unspecified vowel, and then *i/Head and *a/NonHead could determine the vowel quality. This would not then be an instance of the P >> M ranking schema, though it could still be considered a TETU effect.
Another case of SCA that can be analyzed as a TETU effect is found in Turkana (Nilotic, Kenya; Dimmendaal 2000). In this language, the suffix used to form abstract nouns occurs as -isî with CVC roots and -û with CiVjCiVjC roots (examples are from Dimmendaal 2000: 166).

(12) á-rèn-îsî ‘redness’ (< -rèn ‘be red’) á-gógôn-û ‘strength’ (-gógôn ‘be strong’)
    á-jôk-îsî ‘kindness’ (< -jôk ‘be kind’) á-bábár-û ‘saltiness’ (-bábár ‘be salty’)

Since each noun has an á- feminine gender prefix, the resulting forms all have four syllables; this could be driven by foot structure/parsing constraints and could therefore result from TETU.

4. The Subcategorization approach

The existence of non-optimizing SCA demonstrates the need for a mechanism other than output optimization to handle SCA. A subcategorization approach captures the fact that there are cases of SCA where phonological well-formedness considerations have no bearing on the choice among allomorphs in a given environment. Subcategorization has been proposed by Lieber 1980, Kiparsky 1982, Selkirk 1982, Inkelas 1990, Orgun 1996, Yu 2003, inter alii. The main idea, for our purposes, is that the representation of an affix includes requirements for stems to which it will attach.

An example of an SCA pattern that is particularly well-suited to the subcategorization approach is found in Nakanai (Austronesian, New Britain; Johnston 1980). In Nakanai, the -il- form of the nominalizing affix occurs when it can be in the first syllable and adjacent to main stress; -la occurs elsewhere (examples are from Johnston 1980: 177-178).

(13) au ‘steer’ vi-gile-muli ‘tell a story’
    il-au ‘steering’ vigilemulimuli-la ‘story’
    peho ‘die’ vi-kue ‘fight (v.)’
    p-il-eho ‘death’ vikue-la ‘fight (n.)’
    loso ‘dive’ go-ilo ‘go in’
    il-osó ‘diving’ goilo-la ‘entrance’

Stress is on the penult (Johnston 1980: 256). McCarthy (2003: 101-102) claimed that in Nakanai, -il- is attracted to the main stress of the word but is also a ‘formal prefix’, and the pattern is driven by dispreference for a stress shift with respect to the base. In McCarthy’s analysis, OO-PK-MAX penalizes the stress shift caused by -la, while AFX-TO-HD(-il-) and PREFIX/σ(-il-) limit -il- to occurring with disyllabic and smaller stems. Each of these constraints would need to outrank a morphological uniformity constraint (though McCarthy avoids proposing such a constraint). This would therefore be a case of P >> M. Since OO-PK-MAX is not active elsewhere in the language (McCarthy 2003: 102), this is a case of ‘the emergence of the faithful’, not TETU. Thus, even a case of SCA that is apparently phonologically optimizing still contradicts Kager’s (1996) claim that SCA results from TETU.

A subcategorization account for the Nakanai allomorphy is characterized as follows. The -il-affix subcategorizes for the first vowel and main stress (under a cyclic account of stress). -il- can then attach only to disyllabic and smaller stems (see Yu 2003 on using subcategorization for infixation), and -la will attach to all other stems by virtue of its less restrictive subcategorization frame. The Nakanai example is thus exactly the type of case predicted and accounted for by subcategorization.

4.1 Using Subcategorization to model SCA

In this section, I show how subcategorization works using some examples seen above. First, to account for Tzeltal perfective allomorphy, we can propose the subcategorization frames in (14).
(14) Tzeltal perfective construction A
\[
\text{[[\#\sigma\#] verb stem \ oh perf suffix \ perf verb}}\]

Tzeltal perfective construction B (‘elsewhere’)
\[
\text{[[ \verb stem \ eh perf suffix \ perf verb}}\]

In this analysis, the -oh suffix left-subcategorizes for a verb stem with only a single syllable, while the -eh suffix left-subcategorizes for a verb stem with no phonological requirements. This is how -eh is selected as the ‘elsewhere’ allomorph.

In Nakanai, we can characterize the nominalizing allomorphy seen above as subcategorization for a vowel and a stressed syllable. Below are proposed subcategorization frames for the two nominalizing allomorphs in Nakanai.

(15) Nakanai nominalizing const. A
\[
\text{[il nominalizing prefix [V, \sigma] verb stem noun}}\]

Nakanai nominalizing const. B (‘elsewhere’)
\[
\text{[verb stem \ la nominalizing suffix noun}}\]

Here, the il- allomorph right-subcategorizes for the leftmost vowel and stressed syllable of a verb stem. In the event that il- cannot be adjacent to both the leftmost vowel and the stressed syllable, then the elsewhere allomorph, -la, is selected. This allomorph left-subcategorizes for a verb stem with no phonological requirements.

4.2 Estonian revisited

Given the existence of the subcategorization mechanism discussed above, there are two possible reanalyses for the Estonian data that were the impetus for Kager’s claim that SCA is output optimization. The first is the subcategorization account outlined below. Here, I propose subcategorization frames to account for the allomorphy found in the genitive plural (16a) and partitive plural (16b). Each of these is characterized as having the more specific allomorph subcategorize for a stem that ends in a complete foot.

(16)a. Estonian genitive plural const. A
\[
\text{[[Ft\#] stem \ te gen pl suffix \ gen pl word}}\]

B. Estonian genitive plural const. B (‘elsewhere’)
\[
\text{[[ \ stem \ te gen pl suffix \ gen pl word}}\]

b. Estonian partitive plural const. A
\[
\text{[[Ft\#] stem \ sit part pl suffix \ part pl word}}\]

Estonian partitive plural const. B (‘elsewhere’)
\[
\text{[[ \ stem \ it part pl suffix \ part pl word}}\]

A second reanalysis is possible for the Estonian genitive plural and partitive plural: a non-suppletive account involving a single underlying form of the affix for each morphological category and phonological rules specific to each category: a rule of tt degemination for the genitive plural, and a rule of s insertion for the partitive plural. These rules are schematized below.

(17)a. Genitive plural tt degemination
\[
\text{\[\sigma \sigma\]_{Ft}}\]

b. Partitive plural s insertion
\[
\text{\[\sigma \sigma\]_{Ft}}\]

This analysis in terms of phonological rules is possible because the allomorphs in each category are phonologically similar to each other, allowing us to propose a single underlying form (/-tte/ for the genitive plural and /-it/ for the partitive plural). Interestingly, this is true of some other cases of optimizing SCA as well, so that of the 12 cases of optimizing SCA found in this survey, some may not actually involve suppletive allomorphy. For example, in the Shipibo repetitive allomorphy discussed above, it is possible to propose a single underlying form, /-ribV/, that corresponds to the two allomorphs. This is important because not only have we demonstrated that SCA is not always
optimizing, but it may even be the case that the majority of cases of SCA are non-optimizing once we factor out cases of optimizing SCA that do not actually involve suppletive allomorphy.

5. Conclusion

Given the existence of SCA that output optimization cannot handle, and the existence of subcategorization, which can handle these cases as well as those modeled by output optimization, we have 2 options for our theory. Option A is to model the optimizing cases using output optimization and the non-optimizing cases using subcategorization; this is similar to a proposal advanced by Booij (1998). Option B, the option advocated here, is to model all cases of PCSA (including SCA) using subcategorization.

Option B avoids the problem of having multiple theoretical mechanisms to model a single phenomenon. Option B also captures the fact that SCA appears to be a unitary phenomenon, rather than two distinct phenomena that can be easily differentiated, which is what Option A implies. As shown below, cases of SCA are distributed along a continuum of optimization ranging from optimizing, TETU effects (as in Estonian) to anti-optimizing or ‘perverse’ effects (as in Dyirbal).

If we abandon the use of output optimization for PCSA in favor of subcategorization, this constitutes a phenomenon previously analyzed using P >> M for which it is no longer needed. We thus have another argument for the abandonment of P >> M, already advocated by Yu (2003), Blevins (1999), and Paster (in press, in progress). This converges with results for infixation (Yu 2003) and phonologically conditioned affix order (Paster in press), where P >> M is shown to be unnecessary and to make incorrect predictions.

Appendix of SCA examples

Dyirbal (Pama-Nyungan, Australia; Dixon 1972): ergative*
Estonian (Kager 1996): genitive plural, partitive singular, partitive plural
Finnish (Karlsson 1999): illative
Jivaro (Jivaroan, Ecuador; de Maria 1918): genitive*, negative*
Kaititj (Pama-Nyungan, Australia; Koch 1980): ergative/instrumental/locative*
Kashaya (Pomoan, California; Buckley 1992): durative*
Kimagambia (Bantu, Tanzania; Odden 1996): perfective
Nakanai (Austronesian, New Britain; Johnston 1980): nominalizer
Nancowry (Mon-Khmer, Nicobar Islands; Radhakrishnan 1981): causative, instrumental*
Saami (Lappic, Norway; Dolbey 1997): person markers, passive
Shipibo (Panoan, Peru; Elías-Ulloa 2004): ergative, repetitive
Spanish (Harris 1979): diminutive*
Turkana (Nilotic, Kenya; Dimmendaal 2000): abstract noun suffix
Tzeltal (Mayan, Mexico; Walsh Dickey 1999): perfective*
Warlpiri (Pama-Nyungan, Australia; Dixon 1980): ergative*
Zuni (isolate, New Mexico; Newman 1965): singular*

* = allomorph distribution is phonologically unpredictable
References

de Maria, Jose. 1918. Gramatica y Vocabulario Jibaros. Quito: Impr. de la Universidad Central.