A unified constraint-based account of the English indefinite article
Brian W. Smith • brianwsmith@ucla.edu
University of California Los Angeles

Overview and goals

The basic pattern: a before consonants, an before vowels.

<table>
<thead>
<tr>
<th>a</th>
<th>an</th>
</tr>
</thead>
<tbody>
<tr>
<td>a pear</td>
<td>an apple</td>
</tr>
<tr>
<td>a minute</td>
<td>an hour</td>
</tr>
<tr>
<td>a morpheme</td>
<td>an allomorph</td>
</tr>
</tbody>
</table>

Phonologically Conditioned Suppletive Allomorphy (PCSA): there are listed allomorphs; the choice between allomorphs is conditioned by phonology; often one allomorph is the default.

<table>
<thead>
<tr>
<th>Listedness</th>
<th>Phonology</th>
<th>Defaults</th>
</tr>
</thead>
<tbody>
<tr>
<td>/æ/ /ɛ/ /æn/</td>
<td>/æn/ before vowels</td>
<td>/æ/ elsewhere</td>
</tr>
</tbody>
</table>

Three questions about PCSA:

- Is it driven by markedness constraints or lexical subcategorization?
- What is the evidence for defaultness? How to model it?
- Can putative cases of PCSA be reanalyzed with minor rules, or in OT, with morpheme-specific phonology?

Answers from a and an:

- Markedness constraints condition PCSA. Diverse repairs respect "aV":
  - a is an allomorph - [i]-epenthesis; the reduction
  - Default is observed in production, and modeled with UR constraints.
- In OT, an analysis with morpheme-specific phonology is difficult at best.

A is the default

The allomorph a is used in the elsewhere environment. Counts below from the Switchboard Corpus of American English (Godfrey & McDaniel 1992).

Disfluencies. The allomorph a is more likely before disfluencies. This holds regardless of the word following the disfluency.

Before disfluencies, function words are unreduced (Foxtree & Clark 1997).[a] um apple *[a] um apple

Speech errors. 95% of errors in Switchboard are an > a. Phrases like a apple are about 17x more likely than an pear.

<table>
<thead>
<tr>
<th>V-initial N/Adj</th>
<th>C-initial N/Adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>230</td>
</tr>
<tr>
<td>an</td>
<td>4080</td>
</tr>
<tr>
<td>% a</td>
<td>5%</td>
</tr>
</tbody>
</table>

*ə.V across English

*əV Assign one  =$ for every əV sequence in the surface representation.

This constraint has effects across English, driving and blocking a number of different phonological processes. This conspiracy was noticed as early as Stene (1954).

Phonotactics. No hiatus in English where left vowel is lax (Chomsky & Halle 1965) • [ɪə] [ɪɛθ] [ɪʌθ] [ɪʊθ] [ɪʌθ] • 10 counterexamples in the Carnegie Mellon University pronouncing dictionary. • +2,500 examples of hiatus where the left vowel is tense.

[ʔ]-epenthesis. Epenthesis is nearly categorical in the context /æ/ /ɛ/ (Keating et al 1994)

Also within words (Plag 1999)

mora[ʔ] e, samba[ʔ] ing, dada[ʔ] ist

Children up to 10 yrs: a [ʔ] apple (Pak 2014)

Reduction. English speakers are less likely to reduce function words before vowels. (Keating et al 1994).

- [ɜ] apple [ɪ] pear *[ɪθ] apple
- Same results for fo [fʊ] [fʊθ], shown in graph to the right. These data show the proportion of the strong (unreduced) form in the Buckeye corpus of conversational English (Buckeye: Pitt et al 2007). Each box is a speaker in Buckeye.

- Speakers who use prevocalic *[ɪθ] or *[aθ] are still more likely to reduce before consonants.

Account with UR constraints


UR = AN Assign one  =$ for every FS [Indef, Det] that does not correspond to /æn/.

UR = A Assign one  =$ for every FS [Indef, Det] that does not correspond to /æ/.

*əV Assign one  =$ for every əV sequence in the surface representation

REDUCE Assign one  =$ for every function word with an unreduced vowel in a weak prosodic position (based on Selkirk 1972).

Lexically indexed constraints

An alternative to PCSA for a and an: morpheme-specific phonology. This is implemented as a minor rule of n-epenthesis in Rotenberg (1974).

A possible implementation in OT is lexically indexed constraints, either faithfulness (Fukazawa 1997) or markedness (Pater 2010). Indexed markedness misses the conspiracy. Indexed faithfulness can capture both n-epenthesis and [ʔ]-epenthesis, although it cannot capture reduction.

Dep-2. Assign one  =$ for every [ʔ] in the output without an input correspondent, if the [ʔ] follows the morpheme a.

Indexed faith

[əV] = [ə] < [θ] epenthesis to prevent hiatus, not = epenthesis.

D = [ʔ]-epenthesis blocked for a n-epenthesis violation.

No epenthesis before a consonant. REDUCE = IDENT gets reduction.


UR = AN Assign one  =$ for every FS [Indef, Det] that does not correspond to /æn/.

UR = A Assign one  =$ for every FS [Indef, Det] that does not correspond to /æ/.

*əV Assign one  =$ for every əV sequence in the surface representation

REDUCE Assign one  =$ for every function word with an unreduced vowel in a weak prosodic position (based on Selkirk 1972).

Reduction in the introduces a ranking paradox. Why not [ŋ] apple?

In-epenthesis and reduction occur for a, they should also occur for the if reduction is blocked for it, it should also be blocked for a.

The problem: the choice of repair is determined solely by the ranking of constraints. Blocking [ə] and [θ] requires even more indexation. In the PCSA analysis, [əθ] isn't an option because /əθ/ isn't a listed allomorph.

Lexically indexed constraints

An alternative to PCSA for a and an: morpheme-specific phonology. This is implemented as a minor rule of n-epenthesis in Rotenberg (1974).

A possible implementation in OT is lexically indexed constraints, either faithfulness (Fukazawa 1997) or markedness (Pater 2010). Indexed markedness misses the conspiracy. Indexed faithfulness can capture both n-epenthesis and [ʔ]-epenthesis, although it cannot capture reduction.

Dep-2. Assign one  =$ for every [ʔ] in the output without an input correspondent, if the [ʔ] follows the morpheme a.

Indexed faith

[əV] = [ə] < [θ] epenthesis to prevent hiatus, not = epenthesis.

D = [ʔ]-epenthesis blocked for a n-epenthesis violation.

No epenthesis before a consonant. REDUCE = IDENT gets reduction.


UR = AN Assign one  =$ for every FS [Indef, Det] that does not correspond to /æn/.

UR = A Assign one  =$ for every FS [Indef, Det] that does not correspond to /æ/.

*əV Assign one  =$ for every əV sequence in the surface representation

REDUCE Assign one  =$ for every function word with an unreduced vowel in a weak prosodic position (based on Selkirk 1972).