A/an and the: allomorphy or phonology?

Marjorie Pak*

LSA Annual Meeting, Minneapolis, January 2–5, 2014

1. The paradox. The English a/an alternation presents a well-known paradox: it is restricted to a single morpheme, which suggests that it is a morphological phenomenon, yet it depends crucially on information about the following word, and thus cannot be characterized as a strictly word-internal process. Phonological treatments of a/an require lexically restricted, idiosyncratic rules of /n/-insertion or /n/-deletion (Hurford 1972, Perlmutter 1970, Vennemann 1974), while morphological treatments involve relaxing the word-boundedness of morphological operations and introducing some notion of ‘phrasal’ or ‘external’ allomorphy (Hayes 1990, Mascaró 1996).

Mascaró (1996:517) presents an allomorphic treatment of a/an as part of a more general set of claims about external allomorphy as ‘emergence of the unmarked’ (TETU). The idea is that in phrasal allomorphy, allomorph choice is driven by the phonological well-formedness of the surface string rather than by word-internal faithfulness constraints. For example, a and an are listed as suppletive allomorphs of D[-def]; since they are equally faithful, the choice between them comes down to the low-ranked phonological constraints ONSET and NO-CODA.

1. A/an as TETU (Mascaró 1996)

<table>
<thead>
<tr>
<th>(a,an) book</th>
<th>ONSET</th>
<th>No-CODA</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.book</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>an.book</td>
<td>*</td>
<td>**!</td>
</tr>
</tbody>
</table>

A similar analysis has been extended to many other cases of apparent external allomorphy, including French adjectival alternations (e.g. beau/bel), the Korean nominative-case alternation -i/ka, and the Northwest Catalan definite-article alternation lo/l (Mascaró 1996, 2007; Lee 2009).

Here I look at a/an alongside a strikingly similar but less well-known alternation: /ði/~/ðə/ in the English definite article (henceforth THE). The basic pattern (/ði/ before vowels, /ðə/ before consonants) and the underlying syntactic structures are nearly identical to those found with a/an, and given that tense vowels are diphthongized in English, both alternations seem to be phonologically optimizing. We could thus treat THE as another case of allomorphy as TETU.

2. THE as TETU (to be rejected)

<table>
<thead>
<tr>
<th>(ðə, ði) book</th>
<th>ONSET</th>
<th>No-CODA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ðə.book</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>ði.j.book</td>
<td></td>
<td>**!</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(ðə, ði) egg</th>
<th>ONSET</th>
<th>No-CODA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ðə.egg</td>
<td></td>
<td>**!</td>
</tr>
<tr>
<td>ði.j.egg</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Despite these apparent similarities, I argue that THE is not a case of external allomorphy, but is derived by a phonological vowel-reduction rule.

2. A phonological analysis of THE. In deciding whether a given alternation is phonological or allomorphic, we consider (inter alia) (i) how closely the alternants resemble each other phonologically, and (ii) how restricted the alternation is. For example, Korean -i/ka (sok-i

* Thanks to David Embick, Kim Edmunds, Ian Kirby, Chris Naber, and audiences at the 2014 LSA Annual Meeting, the Emory Linguistics Colloquium, SECOL 81, and UPenn F-MART for helpful feedback and discussion. Thanks to Kim Edmunds, Chris Naber and Greg Tracy for help with CHILDES data retrieval and coding, and to the Emory Program in Linguistics for funding for this project. Author: Marjorie Pak, Emory University (mgpak@emory.edu).
This analysis captures the observation that /æn/ and /ən/ are phonologically identical except that one has a full vowel where the other has /ə/; and (ii) V~ə alternations are found in many other contexts in English, with /ə/ typically showing up unstressed and before consonants (as in THE) (Chomsky & Halle 1968:111ff). If we treated THE as suppletive allomorphy, these parallels would have to be viewed as accidental.

3. Vowel reduction also applies in a/an. A/an is often implicitly assumed to be a two-way alternation, but many adults actually have four forms, distributed as in (5) (see also Clark & Fox Tree 2002:102). Notice that the ‘strong’ forms /æŋ/ and /ən/ have full vowels where their ‘weak’ counterparts have /ə/. In other words, a and an behave exactly like other monosyllabic function words that have full vowels when stressed and /ə/ when stressless (3c) – and we can use our VR rule from (4c) to derive the weak forms of both articles (6).

```
(3) a. Stylistic variation: believe, behave, relax, emergency, eraser, enormous, eleven
   b. Word-formation: beaut/i/ ~ beaut/ə/ful, expl/e/n ~ expl/ə/nation
   c. Monosyllabic function words:
      You c/æ/n finish early, but you won’t. ~ You c/ə/n do it.
      John wrote th/i/ ~ John wrote th/ə/páper.

Interestingly, THE may also be easier to acquire than a/an: in our study of 6000+ utterances from the CHILDES corpus (MacWhinney 2000; see References), children’s prevocalic an consistently lags behind prevocalic /ði/ – even though adults in the corpus use prevocalic an more consistently than prevocalic /ði/ (p = 0) (see also Newton & Wells 1999). Conceivably, children are noticing parallels between THE and other V~ə alternations in English and are therefore able to acquire THE more easily than the unprecedented and arbitrary a/an.

Assuming a serialist architecture in which syntactic structures are linearized and spelled out step-by-step in PF (Embick & Noyer 2001), I propose that THE is derived as follows:

```
(4) a. Article Cliticization: English D[±def] is adjoined to the following
   word (a type of Local Dislocation, similar to French Article
   Cliticization in l’arbre, l’école, etc. (Embick 2010:87ff)) [D[±def] [book]]
   b. Vocabulary Insertion: D[±def] is spelled out as /ði/ [ði [book]]
   c. Vowel Reduction (VR): V[-stress] → ə /__C (word-internal) [ðə [book]]

Crucially, VR is a strictly word-internal rule; this explains why it does not apply across word boundaries in carry me, crazy kids, etc. The fact that the behaves like part of the following word for the purposes of VR is accounted for by the Article Cliticization operation in (4a).

3. Vowel reduction also applies in a/an. A/an is often implicitly assumed to be a two-way alternation, but many adults actually have four forms, distributed as in (5) (see also Clark & Fox Tree 2002:102). Notice that the ‘strong’ forms /æŋ/ and /ən/ have full vowels where their ‘weak’ counterparts have /ə/. In other words, a and an behave exactly like other monosyllabic function words that have full vowels when stressed and /ə/ when stressless (3c) – and we can use our VR rule from (4c) to derive the weak forms of both articles (6).

```
(5) a. __V, [+stress]: I want /æn/ apple, not two apples.
   b. __V, [-stress]: I want /ən/ apple.
   d. __{C/ Ø}, [-stress]: I want /ə/ book.

(6) a. Article Cliticization (see (4a)) [D[±def] [book]]
   c. Vowel Reduction (see (4c)) [ə [book]]

This analysis captures the observation that /æŋ/ and /ə/ are similar to each other in the same way as /æŋ/~ /ən/ and /ði/~ /ðə/. At the same time, it captures an important difference between a/an
and THE: While THE is a two-way alternation explained by VR alone, a/an is a four-way alternation involving both allomorphy and VR. Under a uniformly allomorphic treatment, we would have to assume four-way allomorphy for a/an, with spellout rules that insert full-vowel forms when [+stress] and /ə/ variants elsewhere — but leave this correspondence unexplained.

4. Rule-ordering effects: pause-fillers and glottal stops. Under a TETU analysis of a/an and/or THE, where allomorphic choice is driven by the phonological well-formedness of surface strings (1), at least two observations go unexplained. First, the pause-fillers uh and um do not trigger insertion of an, even though they are vowel-initial (Clark & Fox Tree 2002):

(7) a. I’d like {/ej/, ?*an} um... a large coffee and a croissant.
   b. I’d like {*/ej/, an} umbrella.

Second, emphatic glottal stops (frequently inserted at the beginning of a pitch-accented V-initial syllable in English (Garellek 2012)) do not block insertion of an or /ði/, even though /ʔ/ is a consonant. In a TETU-based account, it is unclear how e.g. (8a) would ever surface, given that /a.?idiot/ and /a.n idiot/ are more well-formed phonologically.

(8) a. What an /ʔ/idiot.
   b. That was /ðiʔ/ other guy.

In the current model, we can hypothesize that pause-fillers and emphatic /ʔ/ are not yet present at the stage when Vocabulary Insertion and Vowel Reduction apply, but are added later in PF (cf. Rotenberg’s (1978) treatment of parentheticals); this is why they are ‘invisible’ for the purposes of a/an and THE.

(9) Derivation of /ej/ um...
   a. Article Cliticization (4a): D[-def] (NA; nothing follows D[-def] at this stage)
   b. Vocabulary Insertion: e
   c. Vowel Reduction (4c): e (NA because __C context isn’t met)
   d. Pause-Filler Insertion: e um
   e. Glide Insertion: ej um

(10) Derivation of an /ʔ/idiot
   a. Article Cliticization (4a): [D[-def] [idiot]]
   b. Vocabulary Insertion: [æn [idiot]]
   c. Vowel Reduction (4c): [ən [idiot]]
   d. Emphatic /ʔ/ Insertion: [ən [ʔidiot]]

If this proposal is on the right track, pause-fillers and emphatic /ʔ/ should be visible for rules that apply later in PF, after VR. This prediction is borne out: Flapping, which has been treated as a late rule in other serialist treatments (Kaisse 1985, Seidl 2001, Pak 2008), is conditioned by the initial vowel in uh/um as well as in umbrella, but is blocked by emphatic /ʔ/.

(11) a. I need that/ɾ/ uh ... that/ɾ/ umbrella.

As a late rule, Flapping is dependent on the (final) surface phonetics of an utterance: if /ɾ/ is followed by /ʔ/ on the surface, it cannot become a flap. A/an and THE, while they may play a role in creating optimal syllables, do not operate on surface phonetic strings in the same way as
Flapping. In the current model, this is because their pronunciation is determined by early rules and thus cannot be affected by later PF processes.

References


CHILDES corpora: