

# 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)

{a,an} <b>book</b>	ONSET	NO-CODA
☞ a.book	*	*
an.book	*	**!

{a,an} <b>egg</b>	ONSET	NO-CODA
a.egg	**!	*
☞ a.n egg	*	*

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)

{ðə, ði} <b>book</b>	ONSET	NO-CODA
☞ ðə.book		*
ðij.book		**!

{ðə, ði} <b>egg</b>	ONSET	NO-CODA
ðə.egg	*!	*
☞ ði.j egg		*

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).

‘inside.NOM’ vs. *so-ka* ‘cow.NOM’) is a good candidate for an allomorphic treatment because the alternants (i) bear very little resemblance to each other and (ii) alternate only in the nominative-case particle. English THE, however, is by no means this clear-cut: (i) /ði/ and /ðə/ 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) 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/ə/n finish early, but you won’t. ~ You c/ə/n dó it.  
 John wrote th/í/ paper on parentheticals. ~ 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 /e(j)/ and /æn/ 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.  
 c. \_\_{C/Ø}, [+stress]: I want /e(j)/ book, not two books.  
 d. \_\_{C/Ø}, [-stress]: I want /ə/ book.
- (6) a. *Article Cliticization* (see (4a)) [D[-def] [book]]  
 b. *Vocabulary Insertion*: D[-def] ↔ /æn/ / \_\_V, /e/ elsewhere [e [book]]  
 c. *Vowel Reduction* (see (4c)) [ə [book]]

This analysis captures the observation that /e(j)/ and /ə/ are similar to each other in the same way as /æn/~ /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.ʔidiot/ and /a.n idiot/ are more well-formed phonologically.

- (8) a. What an /ʔ/ídiot.  
b. That was /ði ʔ/óther 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. <i>Article Cliticization</i> (4a): | D[-def] | (NA; nothing follows D[-def] at this stage) |
| b. <i>Vocabulary Insertion</i> :      | e       |   |
| c. <i>Vowel Reduction</i> (4c):       | e       | (NA because __C context isn't met)          |
| d. <i>Pause-Filler Insertion</i> :    | e um    |   |
| e. <i>Glide Insertion</i> :           | ej um   |   |
- (10) Derivation of *an* /ʔ/ídiot
- |                                       |                   |
|---------------------------------------|-------------------|
| a. <i>Article Cliticization</i> (4a): | [D[-def] [idiot]] |
| b. <i>Vocabulary Insertion</i> :      | [æn [idiot]]      |
| c. <i>Vowel Reduction</i> (4c):       | [ən [idiot]]      |
| d. <i>Emphatic /ʔ/ Insertion</i> :    | [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 tha/ɾ/ uh ... tha/ɾ/ umbrella.  
b. That's Fa[t̪, \*ɾ] ʔAlbert, not Flat Stanley. (cf. That's Fa[ɾ] Albert.)

As a late rule, Flapping is dependent on the (final) surface phonetics of an utterance: if /t/ 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

- Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.
- Clark, Herbert H. & Jean E. Fox Tree. 2002. Using *uh* and *um* in spontaneous speaking. *Cognition* 84. 73–111.
- Embick, David. 2010. *Localism versus globalism in morphology and phonology*. Cambridge, MA: MIT Press.
- Embick, David & Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32. 555–595.
- Garellek, Marc. 2012. Glottal stops before word-initial vowels in American English: Distribution and acoustic characteristics. *UCLA Working Papers in Phonetics* 110. 1–23.
- Hayes, Bruce. 1990. Precompiled phrasal phonology. In Sharon Inkelas & Draga Zec (eds.), *The phonology-syntax connection*. 85–108. Chicago: University of Chicago Press.
- Hurford, James R. 1972. The diachronic reordering of phonological rules. *Journal of Linguistics* 8. 293–295.
- Kaisse, Ellen M. 1985. *Connected speech: The interaction of syntax and phonology*. Orlando: Academic Press.
- Lee, Yongsung. 2009. Allomorphy selection: universal and morpheme-specific constraints. In Young-Se Kang et al. (eds.), *Current Issues in Linguistic Interfaces, Volume 2*. 417–434. Seoul: Hankookmunhwasa.
- Mascaró, Joan. 1996. External allomorphy as emergence of the unmarked. In Jacques Durand & Bernard Laks (eds.), *Current trends in phonology: Models and methods*. 473–483. Salford: University of Salford.
- Mascaró, Joan. 2007. External allomorphy and lexical representation. *Linguistic Inquiry* 38(4). 715–735.
- Newton, Caroline & Bill Wells. 1999. The development of between-word processes in the connected speech of children aged between three and seven years. In B. Maassen & P. Groenen (eds.), *Pathologies of Speech and language: Advances in Clinical Phonetics and Linguistics*. 67–75. Whurr Publishers.
- Pak, Marjorie. 2008. *The postsyntactic derivation and its phonological reflexes*. Philadelphia, PA: University of Pennsylvania dissertation.
- Perlmutter, David. 1970. On the article in English. In M. Bierwisch & K. Heidolph (eds.), *Progress in linguistics*. 233–248. The Hague: Mouton.
- Rotenberg, Joel. 1978. *The syntax of phonology*. Cambridge, MA: MIT dissertation.
- Seidl, Amanda. 2001. *Minimal indirect reference: A theory of the syntax-phonology interface*. New York: Routledge.
- Vennemann, Theo. 1974. Restructuring. *Lingua* 33. 137–156.
- CHILDES corpora:*
- Bliss, L. (1988). The development of modals. *The Journal of Applied Developmental Psychology* 9. 253–261.
- Braunwald, S. R. 1993. Differences in two sisters' acquisition of first verbs. ERIC Document Reproduction Service.
- Brown, R. 1973. *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Carterette, E. C., & M. H. Jones. 1974. *Informal speech: Alphabetic and phonemic texts with statistical analyses and tables*. Berkeley, CA: University of California Press.
- Demuth, K., J. Culbertson & J. Alter. 2006. Word-minimality, epenthesis, and coda licensing in the acquisition of English. *Language & Speech* 49. 137–174.
- Garvey, C. 1979. An approach to the study of children's role play. *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition* 12.
- Gathercole, V. 1980. *Birdies like birdseed the bester than buns: A study of relational comparatives and their acquisition*. University of Kansas dissertation.
- Gleason, J. B., & E. Greif. 1983. Men's speech to young children. In B. Thorne, C. Kramerae, & N. Henley (eds.), *Language, Gender and Society*. Rowley, MA: Newbury.
- Kuczaj, S. 1977. The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior* 16. 589–600.
- MacWhinney, B. 2000. *The CHILDES Project: Tools for analyzing talk. Third Edition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Nelson, K. 1989. *Narratives from the crib*. Cambridge, MA: Harvard University Press.
- Sachs, J. 1983. Talking about the there and then: The emergence of displaced reference in parent-child discourse. In K. E. Nelson (ed.), *Children's language, Vol. 4*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Warren-Leubecker, A. 1982. *Sex differences in speech to children*. Unpublished doctoral dissertation. Georgia Institute of Technology.
- Weist, R. M., A. Pawlak & K. Hoffman. 2009. Finiteness systems and lexical aspect in child Polish and English. *Linguistics* 47. 1321–1350.