The acquisition of English article alternations: Variation, competition and the default*

Marjorie Pak
Emory University
mgpak@emory.edu

September 2017
To appear in *GURT 2017: Variable properties: Their nature and acquisition*

Alongside the familiar *a~an* alternation (1a), many English speakers have a similar alternation in the definite article (1b), with [ðə] before consonants and [ði] before vowels:

(1) a. a book, a young man ~ an apple, an old man
b. [ðə] book, [ðə] young man ~ [ði] apple, [ði] old man

Both *a~an* (henceforth A(N)) and [ði~ðə] (henceforth TH(I)) are VARIABLE alternations—meaning specifically that speakers sometimes 'overuse' the preconsonantal forms *a* and [ðə] in prevocalic contexts (%*a apple*, %[ðə] *apple*). The preconsonantal forms are also overused in acquisition: utterances like *I ate a apple* are common until age 6, even among children acquiring the 'standard' pattern in (1).

The focus of this paper is a previously unnoticed contrast between A(N) and TH(I): While both alternations are variable, it appears that TH(I) is *more variable* than A(N) (see §1). For example, many speakers who never say *a apple* sometimes say [ðə] *apple*, and the reverse pattern (✓*a apple*, *[ðə] *apple*) is to my knowledge unattested. Building on earlier work (Pak 2016a), I attribute this contrast to the acquisition of A(N) and TH(I) as fundamentally distinct types of alternations—one allomorphic, one phonological.

(2) PROPOSAL:
a. Unlike A(N), TH(I) is a *phonological* alternation. Specifically, [ði] and [ðə] are derived from a single underlying form [ðɪ] by phonological Tensing and Vowel Reduction—the same rules that are responsible for weak-strong alternations in other English function words, e.g. *to* (§2).
b. When children first recognize TH(I) as a phonological alternation, they can analyze it in more than one way. One option is to simply apply the same Tensing rule to TH(I) that is already in use for *to*—a *stress-sensitive* rule for the speakers studied here (see §3). Another option is to adopt a slightly different Tensing rule for TH(I), recognizing that [i] can be tensed even when stressless (*happy, baby*, etc.). The coexistence of these two options—*both* viable strategies for unifying TH(I) with preexisting phonological patterns—is what causes TH(I) to be more variable than A(N). A(N), in contrast, is recognized from the outset as idiosyncratic and thus analyzed allomorphically; there is never a drive to unify it with 'other' phonological patterns.

* Thanks to Ian Kirby, Orli Hendler, the Emory Program in Linguistics and the audience at GURT 2017 for helpful discussion and support. Any errors are my own.

1. The pattern: Variation in A(N) and TH(I)

There is an extensive body of previous work on variation in both A(N) and TH(I), including sociolinguistic studies (Ash and Myhill 1986; Britain and Fox 2009; Gabrielatos et al. 2010), corpus-based studies (Jurafsky et al. 1998; Keating et al. 1994; Pak 2016a; Todaka 1992), and experimental studies (Gaskell et al 2003; Newton and Wells 1999; Raymond, Fisher, and Healy 2002; Raymond et al. 2009). Again, VARIATION is defined here—as in most previous work—as the apparent overuse of the preconsonantal forms *a* and [ðə] in prevocalic contexts:

(3) a. % a apple, % a old man
b. % [ðə] apple, % [ðə] old man

A glottal stop [ʔ] can be inserted between the article and the following word in (3) (*a ʔapple*). Glottal stops are also frequently inserted after the prevocalic forms *an* and [ði] (*I want an ʔapple*) (Pak 2016b). I will not include glottal stops in examples here since the current analysis is not affected by their presence or absence.

1.1 The adult pattern: More variation in TH(I)

Prevocalic *a* is common in many 'non-standard' varieties of English, with reported frequencies up to 100% in Philadelphia AAVE (Ash and Myhill 1986), around 75% among Bangladeshi adolescents in London Towers Hamlets (Britain and Fox 2009) and at somewhat lower levels in London Hackney and Havering (Gabrielatos et al. 2010). Lass (2002) also identifies prevocalic *a* as a feature of the 'extreme' variety of South African English. See Gabrielatos et al. (2009) for a more detailed review of A(N) variation.

Prevocalic [ðə] is attested in 'standard' as well as 'non-standard' varieties of English—e.g. the Buckeye corpus of Central Ohio speech (Raymond et al. 2009, 108), the Switchboard corpus of American English (Jurafsky et al. 1998, 3113), and the TIMIT corpus of read sentences in American English (Keating et al. 1994). This last study reports no prevocalic [ðə] in speakers over 50, but 33% prevocalic [ðə] in younger speakers, suggestive of a change in progress.

Those studies that have looked at A(N) and TH(I) in tandem have shown either high variation in both alternations (e.g. Ash & Myhill 1986), or higher variation in TH(I) than in A(N). For example, Lass (2002) describes middle-class South African English as having variable prevocalic [ðə] but no prevocalic *a*. Other studies showing this asymmetry are summarized in Table 1. Notably, *no* study to my knowledge has found variable A(N) with invariant TH(I).

Table 1. Previous studies of variation in A(N) and TH(I)

	Prevocalic [ðə]	Prevocalic *a*
NA-English speaking adults in CHILDES[a]	10%	4%
White adolescent boys in London Towers Hamlets[b]	35%	25%
Bangladeshi adolescent boys in London Towers Hamlets[b]	81%	75%
Buckeye corpus[c]	~40%	<8%

[a] Pak (2016a, Table 1) [b]Britain and Fox (2009, 189-190) [c]Raymond et al. (2002, 629)

1.2 The pattern in children: More overuse errors with A(N)

As noted in the introduction, young English-speaking children frequently 'overuse' the preconsonantal forms *a* and [ðə] (Gaskell et al. 2003; Seliger 1979):

(4) a. Boba Fett is a animal. (3;11) (MacWhinney 2000, 47a1)
b. I'm gonna buy all [ðə] action figures. (4;02) (MacWhinney 2000, 48b2)

While the utterances in (4) are superficially parallel to those in (3), this is not a simple mirroring of adult variation. Children actually overuse *a more* than they overuse [ðə]—in striking contrast to the adult pattern in Table 1.

Newton and Wells (1999), in an experiment eliciting A(N) and TH(I) in British 3- to 7-year-olds, find the frequency of *an* (over all prevocalic *a(n))* lagging behind the frequency of [ði] (over all prevocalic *the*) by 5 to 15 percentage points at every age. I identify a similar pattern in my CHILDES study of North American English (Pak 2016a, Table 2)—prevocalic *an* lagging behind prevocalic [ði] throughout early childhood—even though the adults in the same corpus had the reverse pattern. This contrast gives rise to an apparent paradox:

(5) Children seem to take longer to acquire prevocalic *an* than prevocalic [ði], even when their caregivers are more consistent in using prevocalic *an* than prevocalic [ði].

Table 2 shows this same contrast borne out within two individual families from CHILDES: MacWhinney (2002) and Braunwald (1993), a subset of the corpora studied in (Pak 2016a). Both of these corpora have a span of at least three years of recorded naturalistic speech produced by children ages 2-7, plus copious speech produced by parents. Moreover, the parents in these two corpora have remarkably similar patterns with A(N) and TH(I) as well as with a third alternation, TO; this will be a point of focus in §3. For now, the key points to be taken from Table 2 are:

- The parents in both families have near-categorical prevocalic *an*, but allow some variation with TH(I) (approximately 15% prevocalic [ðə]).
- All four of the children 'overuse' prevocalic *a* and [ðə] to at least some degree.
- With TH(I), three of the four children show a clear developmental trend—increasing use of prevocalic [ði] with increasing age—and both older siblings use prevocalic [ði] the majority of the time by the end of the study periods.
- With A(N), only one of the four children (MacWhinney Child 1) reaches >50% prevocalic *an* by the end of the study period. Assuming that the other three children did eventually acquire their parents' pattern, any developmental trend they followed must have been established later than the ages studied here.

In §2 I explain the contrast in (5) by proposing that A(N) and TH(I) are fundamentally distinct types of alternations: A(N) is allomorphic while TH(I) is phonological. (The analysis in §§2.2-2.3 is taken more or less directly from Pak 2016a.) In §3 I return to the observation we started with—that among adults, TH(I) is more variable than A(N)—and show that a phonological treatment of TH(I) helps us explain this contrast as well.

Table 2. Frequency of prevocalic *an* and [ði] over a 3- to 5-year period in two families

	speaker	age	% an	an/(a+an)	% ði	ði/(ðə+ði)
Braunwald	**parents**		**100%**	106/106	**84.9%**	152/179
	child 1	4-5	36%	5/14	21%	5/24
		6-7	31%	5/16	90%	53/59
	child 2	2-3	12%	3/25	36%	13/36
		4-5	20%	1/5	38%	3/8
MacWhinney	**parents**		**99%**	238/240	**84%**	338/401
	child 1	3	40%	8/20	38%	18/47
		4-5	29%	9/31	48%	35/73
		6-7	93%	28/30	66%	37/56
	child 2	3	33%	5/15	24%	5/21
		4-5	21%	8/37	38%	50/131

2. A(N) is harder because it is unprecedented

2.1 Stages of acquisition

Let us assume that, in order to reach adultlike patterns with A(N) and TH(I), children must proceed through three steps:

(6) i. Recognizing the alternation (i.e. associating the variants as 'two ways of saying the same word')
ii. Analyzing the alternation (e.g. as allomorphic or phonological)
iii. Consistently executing the alternation in production

Achieving even this first step is likely to be delayed with A(N) and TH(I), due to the way the definite and indefinite articles are distributed in discourse.

First, the distribution is skewed: the ratio of preconsonantal to prevocalic articles is 9:1 in the MacWhinney and Braunwald corpora, meaning that children hear far more preconsonantal forms (*a* and [ðə]) than prevocalic forms (*an* and [ði]) in the input.

Second, the articles are almost always deaccented. English syntax does not allow articles in positions where they would be assigned phrase-final stress, and while it is possible for articles to have contrastive or 'focal' stress (7), such utterances seem to be infrequent in child-directed speech (I have yet to encounter one in my CHILDES corpus work).

(7) a. Mary is thé expert on bats.
b. I said I wanted á car, not two cars.

Consider, by way of contrast, the English preposition and infinitive marker *to*, which participates in a similar alternation [tu~tə] (henceforth TO). Unlike the articles, *to can* be phrase-final, and in this position it is always assigned stress and pronounced with a full vowel:

(8) a. I told John not [tə] go, but he really wants [tù].
b. He gave this present [tə] Sam; I don't know who he gave that one [tù].

Correspondingly, as we will see in §3 (Table 3), the MacWhinney and Braunwald children acquire adultlike TO by age 3, with *no* initial stage of 'overusing' [tə]. This means that the delay in acquiring A(N) and TH(I) is not due to an inability to analyze or control alternations ((6)-ii or (6)-iii). More likely, the culprit is a failure to even notice that there *is* an alternation ((6)-i), because the evidence is not phonologically salient enough.

I therefore assume that prior to step (6)-i, children simply have *a* and [ðə] as *non-alternating* forms of the indefinite article (D[-def]) and definite article (D[+def]), respectively (cf. Gaskell et al. 2003).

(9) Non-alternating *a* and [ðə]
a. D[-def] ↔ ə
b. D[+def] ↔ ðə

Acquiring adultlike A(N) and TH(I), then, involves 'unlearning' the insertion rules in (9) and replacing them with rules that produce the relevant alternations. The crucial difference between A(N) and TH(I) arises at step (6)-ii (analyzing the alternation): Children analyze A(N) allomorphically, TH(I) phonologically.

2.2 Analysis of A(N)

With A(N), there appears to be an additional delay in reaching step (6)-ii. Newton and Wells (1999, 72) show that even when repeating a short sentence like *Jane gave me an ice cream* directly after an adult, children fail to produce *an* 40% of the time (cf. 30% with [ði]). An example from MacWhinney (2002, 20a2) illustrates the same kind of error:

(10) Father: And then he said, "That's an elephant."
Child (2;6): "That's a elephant."

Such imitation errors, reminiscent of the famous *Nobody don't likes me* parent-child exchange from MacNeill (1966, 69), can be interpreted as suggested there: 'Children assimilate… adult models to their current grammars' and do not produce forms that are not part of their grammar, even in imitation.

The reason for this delay in analyzing A(N), as suggested by Pak (2016a), is that A(N) it is unprecedented: this is the only context where [n] alternates with Ø in English. The only solution available to children is to learn A(N) by 'brute force': memorizing and storing the two variants and their contexts as a case of suppletive allomorphy (see Joseph 1997; Kaisse 1985; Rotenberg 1978; among others, for precedent).

(11) D[-def] ↔ ən /__V
↔ ə elsewhere

A [ə] is identified as the 'elsewhere'allomorph because it is selected in pre-pausal and isolation contexts (Rotenberg 1978), e.g. *Let's call it a, a, a… urf* (MacWhinney 2002, 47ba). See Pak (2016a) for an account of other variants of the indefinite article, [ej] and [æn/ɛn].

An additional assumption I make is that allomorphy, normally a 'word-internal' phenomenon, is enabled here by cliticization of the article onto the first word in its complement:

(12) Article Local Dislocation: [D[±def]] ⁀ [X] → [D[±def] [X …]]

(See Embick 2008 for more on Local Dislocation.)

The final stage of acquiring A(N) (step (6)-iii) can be characterized as practicing—remembering to execute (11) instead of the earlier, non-alternating (9)a—until the target frequency is reached. In some dialects, (9)a and (11) continue to coexist (or 'compete') through adulthood; these are dialects that allow *a apple* as well as *an apple* (§1.1). (See Embick (2008) and Kroch (1994) for more on competing grammars.) In other dialects, the non-alternating grammar (9)a becomes extinct, yielding categorical prevocalic *a*; this is presumably the case for the MacWhinney and Braunwald adults in Table 2.

2.3 Acquisition of TH(I)

Despite its apparent similarity to A(N), I argue that TH(I) is a fundamentally different kind of alternation—phonological rather than allomorphic (Pak 2016a). Unlike A(N), TH(I) is *not* unprecedented: there are other contexts where full vowels alternate with [ə] in English, including TO (Chomsky & Halle 1968, 111ff; Jurafsky et al. 1998; Selkirk 1995).

(13) a. *Affixation:* beaut[i]~beaut[ə]ful, happ[i]~happ[ə]ly, etc.
b. *Function words:* tu~tə, kæn~kən, fɔr~fɹ̩, etc.

The parallel between the V~ə alternations in TH(I) and (13) helps explain why TH(I) is acquired earlier than A(N). With A(N), children are forced to create an *ad hoc* allomorphy rule from scratch. TH(I), by contrast, can be subsumed under a more general pattern that has *already been analyzed*.

Specifically, V~ ə alternations can be attributed to phonological rules of *Tensing* and *Vowel Reduction*—'word-internal' rules adapted from Chomsky and Halle (1968, 111ff) that are enabled here by article cliticization (12). (The word-boundedness of Tensing and Vowel Reduction explains why e.g. *ver*[i] *interesting* does not alternate with *ver*[ə] *funny*.) Under this analysis, [ði] and [ðə] are derived from the same underlying form [ðɪ] (14)a. If [ðɪ] is prevocalic or non-cliticized, Tensing applies, yielding [ði]; otherwise, Vowel Reduction applies, yielding [ðə] (14)b.

(14)

		_ apple	_ book
a.	D[+def] ↔ ðɪ	ðɪ apple	ðɪ book
b.	Tensing: V[-low] → [+tense] / __{V,#}	ði apple	--------
	Vowel Reduction: V[-stress -tense] → ə	-------	ðə book

As with A(N), the rest of acquisition (step (6)-iii) involves practicing—remembering to execute grammar (14) (with underlying [ðɪ] and phonological Tensing/Reduction) instead of the earlier grammar in (9)b (with non-alternating [ðə])—until the target frequencies are reached for the given variety of English. Again, I use competing grammars to model adult variation: the Braunwald and MacWhinney parents, for example, retain grammar (9)b alongside grammar (14) and continue to employ it about 15% of the time.

3. TH(I) is more prone to variation because it *is* precedented

We are now ready to address the question we started with: Why do adults have *more* variation in TH(I) than in A(N)? There is no *a priori* reason to expect this contrast, given that variation is widely attested in both phonology and morphology (see Tamminga, MacKenzie, and Embick 2016, 319ff). Still, I will argue that TH(I) is more prone to variation than A(N) due in part to its phonological status—or more specifically, its 'precedentedness.' I propose that different speakers adopt different strategies in their quest to unify TH(I) with other V~ə alternations, e.g. TO.

TO has been described as very similar to TH(I), with [tu] appearing before vowels and [tə] before consonants (Britain and Fox 2009; Ladefoged and Johnson 2015, 118).

(15) a. John wants [tə] see me. ~ John wants [tu] invite me.
b. I went [tə] Dallas. ~ I went [tu] Atlanta.

The MacWhinney and Braunwald adults, however, have a different pattern for TO than for TH(I): they do not show a preference for [tu] in *all* prevocalic contexts, but only if the following vowel is unstressed (e.g. *to Atlánta, to arríve*). If the following word starts with a stressed vowel (e.g. *to ánimals, to éat*), they prefer [tə]. A series of examples produced by the MacWhinney father illustrates this general pattern:

(16)

<u>before V[-stress]</u>	<u>before V[+stress]</u>
you can count [tu] a thóusand? (62a1)	Then you can count [tə] éighty-five. (62a1)
It's just hard [tu] achíeve. (61b1)	Is it okay [tə] éat cookies? (60b1)

Moreover, as noted in §2.1, the MacWhinney and Braunwald children have the same TO pattern as their parents—unlike with TH(I), there is *no* initial stage of 'overusing' [tə].

Table 3. Prevocalic [tu] (vs. [tə]) in MacWhinney and Braunwald

		Adults		*Children*	
		% tu	tu/(tu+ tə)	% tu	tu/(tu+ tə)
Braunwald	before V[-stress]	88%	22/25	71%	15/21
	before V[+stress]	16%	20/122	18%	11/63
MacWhinney	before V[-stress]	83%	68/82	96%	45/47
	before V[+stress]	11%	24/210	20%	25/126

I attribute the pattern in Table 3 to the following *stress-sensitive* Tensing rule, which requires the underlying lax vowel in *to* to have at least some stress in order to be tensed:

(17) $V[-low] \rightarrow [+tense] \; / \; \underline{\quad\quad} \; \{V, \#\}$
$[+stress]$

If the vowel has no stress, or if it is preconsonantal, it will not undergo Tensing and will later be subjected to Vowel Reduction (as in (14)).

This treatment explains why these speakers so often have [tu] before unstressed vowels (*to* is assigned stress here, to avoid a lapse) and [tə] before stressed vowels (*to* doesn't get stressed

here, to avoid a clash). The apparent variability of Tensing by this hypothesis is actually variability in stress assignment.

(18) a. he really wants [tù] (100% [tu]; obligatory stress on *to*)
b. cóunt [tù] a thóusand (85% [tu]; stress lapse in the other 15%)
c. cóunt [tə] éighty-fíve (85% [tə]; stress clash in the other 15%)

Now, assuming that the MacWhinney and Braunwald children have acquired a rule like (17) by age 3, why don't they simply deploy this same rule when they are ready to analyze TH(I)?

There may in fact be speakers who do exactly this. Participants in Raymond et al. (2009, 97-98) more frequently showed a preference for prevocalic [ðə] if the following word had first-syllable stress ([ðə] *ápple*) than if it had second-syllable stress ([ðə] *América*n)—as we would expect if the stress-sensitive rule in (17) were simply extended to TH(I).

The MacWhinney and Braunwald speakers, however, seem to have postulated a slightly different, *non*-stress-sensitive tensing rule for TH(I) (revised from (14)b).

(19) ɪ → [+tense] / __{V,#}

The precedent for this distinction, I believe, comes from the unique status of [i] in many varieties of English. Unlike other tense vowels in English, [i] frequently bears no stress at all.

(20) happy, funny, silly, baby, coffee, copy, Annie, very, lovely, etc…

The stresslessness of [i] in words like (20) can be established by the absence of aspiration on a preceding voiceless stop (21)a, the obligatory flapping (in American English) of a preceding alveolar stop (21)b, and the placement of the mid-tone in the vocative chant on the *final* syllable (in words with secondary stress somewhere before the final syllable, the drop to M-tone goes there instead) (21)c:

(21)	a.	*Aspiration:*	có[p]y, snár[k]y	*cf*. snów[p^{h}]èa, bar[k^{h}]èeper
	b.	*Flapping*:	sée[ɾ]y ($^{??}$see[d]y)	*cf*. thrée-[dì] glásses
	c.	*Vocative chant*:	snárkier, cópier	*cf*. bárkèeper, síghtsèer
			H H **M** H H **M**	H **M** M H **M** M

So there is abundant evidence that [i] can be simultaneously [+tense] and [-stress] in English—unlike [u], for which there is little evidence (Chomsky and Halle 1968 list *voodoo, Hindu* and *jujitsu* as isolated examples of stressless word-final [u]). The unique ability of [i] to be [+tense -stress] is, by hypothesis, why the MacWhinney and Braunwald speakers have analyzed a slightly different Tensing rule for TH(I) (19) than for TO (17).

Summarizing the proposal so far: At the point when children are ready to analyze TH(I), they have two options. Option (i) (possibly adopted by some of Raymond et al.'s participants) is to subsume TH(I) under the existing stress-sensitive Tensing rule for TO, while Option (ii) (adopted by the MacWhinney and Braunwald speakers) is to postulate a slightly different Tensing rule for TH(I) in recognition of the unique status of [i] in English phonology. Both analyses are grounded in aligning TH(I) with some pre-established phenomenon, and thus both follow the spirit of the current proposal, where TH(I) is analyzed phonologically because it is precedented.

(22) Option i: Same (stress-sensitive) Tensing rule for TH(I) and TO

V[-low] → [+tense] / _____ {V, #}
[+stress]

[ðə] cat ~ [ðə] ápple ~ [ði] umbrélla, [tə] go ~ [tə] éat~ [tù] arríve

Option ii: Different Tensing rules for TH(I) and TO

ɪ → [+tense] / ____ {V, #} [ðə] cat ~ [ði] apple ~ [ði] umbrella

ʊ → [+tense] / _____ {V, #} [tə] go ~ [tə] éat~ [tù] arríve
[+stress]

There is in fact a third logical possibility—treating both TO and TH(I) as *non*-stress-sensitive. This would result in a pattern where *to* is pronounced [tu] before *all* vowel-initial words, regardless of stress:

(23) Option iii: Same (non-stress-sensitive) Tensing rule for TH(I) and TO

V[-low] → [+tense] / ____ {V, #}

[ðə] cat ~ [ði] apple ~ [ði] umbrella, [tə] go ~ [tu] eat~ [tu] arrive

This option is also attested: Britain and Fox (2009, 186-187) report >90% prevocalic [ði] *and* [tu] ([ði] *apple* as well as [tu] *Ipswich*) in Fens English.

I believe that the situation seen here—where there are multiple options that all enable speakers to unify TH(I) with some preexisting pattern—is the source of the increased variation we see with TH(I) compared to A(N). Notice that one of these options—Option (i)—will consistently yield *more* prevocalic [ðə] than the others. In a community where some speakers adopt Option (i) while others adopt Options (ii) or (iii), the non-alternating grammar from early childhood will end up being reinforced, thus perpetuating prevocalic [ðə] among all speakers.

To see how this works, imagine two speakers S1 and S2. Both have the early-childhood [ðə]-inserting grammar (24)a (repeated from (9)b) coexisting with a grammar where [ði] and [ðə] are derived phonologically from [ðɪ] (24)b. But they have slightly different versions of (24)b: S1 has gone with Option (i) above, with stress-sensitive Tensing, while S2 has gone with Option (ii), Tensing before all vowels regardless of stress:

(24) a. D[+def] ↔ ðə (non-alternating grammar)

b. S1: V[-low] → [+tense] / _____ {V, #}
[+stress]

S2: ɪ → [+tense] / ____ {V, #}

Now suppose that S1, using their grammar (b), says *the apple*. *The* is not assigned stress in this context (to avoid a clash with *ápple*), and so Tensing does not apply. Subsequently, the vowel in *the* is reduced: [ðə] *apple*.

S2 cannot analyze [ðə] *apple* with their grammar (b); their Tensing rule would apply here, yielding [ði] *apple*. S2 can only analyze [ðə] *apple* with their *non*-alternating grammar (a). S1's utterance thus serves to reinforce S2's non-alternating grammar (a) (cf. Legate and Yang 2007).

Having (re-)confirmed that grammar (a) is viable, S2 later uses it to produce a sentence where [ðə] is pitch-accented: *Mary is* [ðə́] *expert on bats*. Since this utterance can *only* be generated by the non-alternating grammar (a) (none of the Tensing rules in Options (i)-(iii) would fail to apply here), grammar (a) will again be reinforced. And so on.

With A(N), this kind of situation would never arise. There are of course speakers with variable prevocalic *an* (§1.1), and I explained this variation by adopting the same basic assumption as for TH(I): these speakers have a non-alternating (*a*-inserting) grammar coexisting with an alternating (*a~an)* grammar (§2.2). The crucial difference between A(N) and TH(I) is that there is *only one approach* to analyzing the alternating *a~an* grammar: creating an *ad hoc* rule from scratch, because A(N) is recognized from the outset as idiosyncratic. The type of scenario just laid out, where the non-alternating grammar is indirectly reinforced by misanalysis of an alternating grammar, would only arise with alternations with more than one viable analysis.

4. Summary and conclusion

The fact that A(N) and TH(I) both apply to articles, and under such similar conditions, provides an unusual testing ground for hypotheses about the morphology-phonology interface. The focus of this paper was: Why do we find more variation in TH(I) than in A(N)? While variation in allomorphy and phonology are both attested, this paper proposes a reason why some phonological alternations may be more susceptible to variation than parallel allomorphic alternations.

Following on earlier work, I argued that while A(N) and TH(I) are quite similar at first sight, they have different analyses: A(N) is allomorphic; TH(I) is phonological. I then proposed that phonological nature of TH(I) makes it more prone to variation: If an alternation is phonological, it will be analyzed as part of some broader general pattern in the language, but it is not always clear which general pattern it is to be aligned with. In the case of TH(I), we saw three viable strategies that speakers could (and apparently do) adopt: two that treat TH(I) on par with TO, one that recognizes the unique distribution of [i] in English. This kind of indeterminacy, I suggested, allows inter- and intraspeaker variation to perpetuate.

References

Ash, Sharon, and John Myhill. 1986. “Linguistic Correlates of Inter-Ethnic Contact.” In *Current Issues in Linguisitc Theory: Diversity and Diachrony*, edited by David Sankoff. Philadelphia: John Benjamins.

Braunwald, Susan R. 1993. “Differences in Two Sisters’ Acquisition of First Verbs.” ERIC, ED 358 935.

Britain, David, and Sue Fox. 2009. “The Regularisation of Hiatus Resolution in British English.” In *Vernacular Universals and Language Contacts*, edited by Markku Filppula, Juhani Klemola, and Heli Paulasto, 177–205. New York: Routledge.

Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.

Embick, David. 2007. “Linearization and Local Dislocation: Derivational Mechanics and Interactions.” *Linguistic Analysis* 33:303–336

Embick, David. 2008. “Variation and Morphosyntactic Theory: Competition Fractionated.” *Language and Linguistics Compass* 2(1):59–78.

Gabrielatos, Costas, Eivind Nessa Torgersen, Sebastian Hoffmann, and Susan Fox. 2010. “A Corpus-Based Sociolinguistic Study of Indefinite Article Forms in London English.” *Journal of English Linguistics* 38(4):297–334.

Gaskell, M. Gareth, Helen Cox, Katherine Foley, Helen Grieve, and Rachel O’Brien. 2003. “Constraints on Definite Article Alternation in Speech Production: To ‘Thee’ or Not to ‘Thee’?” *Memory & Cognition* 31:715-727.

Joseph, Bryan. 1997. “On the linguistics of marginality: The centrality of the periphery.” In *CLS 33: Papers from the Main Session*, 197–213. Chicago: Chicago Linguistic Society.

Jurafsky, Daniel, Alan Bell, Eric Fosler-Lussier, Cynthia Girand, and William Raymond. 1998. "Reduction of English Function Words in Switchboard." In *Proceedings of the 5th International Conference on Spoken Language Processing*, 3111–3114. Sydney.

Kaisse, Ellen M. 1985. *Connected Speech: The Interaction of Syntax and Phonology*. Orlando: Academic Press.

Keating, Patricia A., Dani Byrd, Edward Flemming, and Yuichi Todaka. 1994. "Phonetic Analysis of Word and Segment Variation Using the TIMIT Corpus of American English." *Speech Communication* 14:131–142.

Kroch, Anthony. 1994. "Morphosyntactic Variation." In *Papers from the 30th Regional Meeting of the Chicago Linguistic Society, Vol. 2*, 180–201. Chicago: Chicago Linguistic Society.

Ladefoged, Peter, and Mark Johnson. 2015. *A Course in Phonetics*. Stamford: Cengage Learning.

Lass, Roger. 2002. "South African English." In *Language in South Africa*, edited by Rajend Mesthrie, 104-126. Cambridge: Cambridge University Press.

Legate, Julie Anne, and Charles Yang. 2007. "Morphosyntactic Learning and the Development of Tense." *Language Acquisition* 14(3):315-344.

MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk*, 3rd edition. Mahwah: Lawrence Erlbaum.

McNeill, David. 1966. "Developmental Psycholinguistics." In *The Genesis of Language: A Psycholinguistic Approach*, edited by Frank Smith and George A. Miller, 15-84. Cambridge: MIT Press.

Newton, Caroline, and Bill Wells. 1999. "The Development of Between-Word Processes in the Connected Speech of Children Aged Between Three and Seven Years." In *Pathologies of Speech and Language*, edited by Ben Maassen and Paul Groenen, 67–75. London: Whurr Publishers.

Pak, Marjorie. 2016a. "How Allomorphic Is English Article Allomorphy?" *Glossa*, 1(1):20.1–27.

Pak, Marjorie. 2016b. "Optimizing by Accident: *A/an* Allomorphy and Glottal Stop." In *Proceedings of the Linguistic Society of America Annual Meeting*, 1, 12:1-13.

Raymond, William D., Julia A. Fisher, and Alice F. Healy. 2002. "Linguistic Knowledge and Language Performance in English Article Variant Performance." *Language and Cognitive Processes* 17:613–662.

Raymond, William D., Alice F. Healy, Samantha McDonnel, and Charlotte A. Healy. 2009. "Acquisition of Morphological Variation: The Case of the English Definite Article." *Language and Cognitive Processes* 24:89-119.

Rotenberg, Joel. 1978. "The Syntax of Phonology." Ph.D. Dissertation, MIT.

Seliger, Herbert W. 1979. "On the Nature and Function of Language Rules in Language Teaching." *TESOL Quarterly* 13:359-369.

Selkirk, Elisabeth. 1995. "The Prosodic Structure of Function Words." In *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, edited by James L. Morgan and Katherine Demuth, 187–214. New York: Lawrence Erlbaum.

Tamminga, Meredith, Laurel MacKenzie, and David Embick. 2016. "The Dynamics of Variation in Individuals." *Linguistic Variation* 16(2):300-336.

Todaka, Yuichi. 1992. "Phonetic Variants of the Determiner 'the.'" In *UCLA Working Papers in Phonetics* 81, 39–47.