

**The threshold of productivity and the “irregularization” of verbs
in Early Modern English¹**

Don Ringe & Charles Yang, U. of Pennsylvania

[Running head: **Threshold of productivity and “irregularization” of verbs**]

¹ We adopt the definition of the *Penn Parsed Corpus of Early Modern English (PPCEME)*, which includes documents from about 1500 to 1710 or so. We are grateful to three anonymous reviewers for helpful suggestions of many kinds; all remaining errors, shortcomings and infelicities are ours.

Abstract.

This paper attempts to determine whether Yang's Tolerance Principle (TP) is useful in explaining one type of morphological change in the history of English. We suggest that some, but not all, innovative non-default past tenses that appear in the written record between 1500 and 1700 can be explained by the TP and thus could be the results of innovation in native language acquisition (NLA).

Key words.

Tolerance Principle, native language acquisition, Early Modern English, irregularization.

1. The Tolerance Principle, its range and limits.

Yang 2016: 8–9, 60–6 proposes and argues for a threshold for the productivity of linguistic rules of all kinds in native language acquisition (NLA), as follows. Let N be the number of lexemes which meet the structural description of the rule and e the number of exceptions (i.e., the number of lexemes which meet the structural description but are not subject to the rule). There is then a threshold

$$\theta_N = \frac{N}{\ln N}$$

such that e must be less than or equal to θ_N for the rule to be productive; that is, the threshold is the number of lexemes that could obey the rule divided by the natural logarithm of the same. If that threshold is not exceeded, so that the rule is productive, the native learner will extend the range of the rule's application, removing exceptions one by one. If the threshold of productivity is exceeded by the number of exceptions, no change is expected; if change is observed, it must proceed by some other mechanism.

This “Tolerance Principle” (TP), specifying the maximum number of exceptions which a rule can tolerate in order to be productive, is compatible with a wide range of phenomena observed in NLA, including the formation of the past tense in English (Yang 2016: 26–34, 81–91), various details of verb inflection in German and Spanish (*ibid.* pp. 34–7, 80–1), and the scope of syntactic constructions. In this paper we focus on

the productivity of lexical classes, i.e. arbitrary inflectional classes of major words which compete for membership, in NLA and otherwise.

Though the TP makes a prediction about any rule a native learner can come up with, even a “crazy” rule, it is not a hypothesis about the origins of all linguistic innovations. It cannot assess an innovation which is not rule-governed, and it has nothing to say about the innovative behavior of adult native speakers. Interestingly, there is some evidence that innovative non-default past tenses in English might be constructed by adults. An especially clear example relevant to the present study emerged from the celebrated “wug” study (Berko 1958). Berko presented both children and adults with nonce verbs that strongly resembled existing English “irregulars”, e.g. “gling”, potentially of the type *ring, rang* or *sling, slung*, and “spow”, potentially of the type *grow, grew*. The past tenses for these verbs produced by children were almost exclusively default pasts in *-ed*, but a significant number of adults suggested *glang* or *glung*. (See Yang 2016: 31–40 for a crosslinguistic review of how children’s acquisition of morphology reflects a categorical distinction between productive and unproductive processes; this is just one example of many.) This dichotomy between changes in the scope of rules in NLA, governed by the TP, and other processes of change will turn out to be important in our investigation. We will return briefly to adult native speaker innovations at the end of section 9.

We also wish to emphasize that this line of work is fundamentally different from earlier approaches, in that it hypothesizes a specific PROCESS in the course of NLA that gives rise to some, but not all, linguistic changes. It contrasts sharply with the approach of Strik 2015, for example, who treats “analogy” as a kind of black box and attempts to determine what is going on inside using purely statistical methods.² That and the (extensive) other work investigating frequency effects in linguistic change is not directly relevant to our inquiry because it does not deal with the TP, nor with NLA at all. In short, we here pursue a new line of inquiry based on radically shifted premises. For a more extensive demonstration of the explanatory power of the TP in dealing with the linguistic past see now Kodner 2020.

² In fact some of Strik’s closing observations, which apparently puzzled him, point to NLA as a source of change. Overall resistance to morphological change (Strik 2015: 183) is a consequence of the fact that children are remarkably good at learning complex inflectional systems; the fact that both very frequent lexemes and rare lexemes tend to resist change (ibid. p. 184) results from the fact that the former are constantly reinforced, while many of the latter are not acquired early enough in NLA (see further section 5 below).

2. English verbs in *-ing*.

An especially clear illustration of the TP is furnished by the past tense of (standard) English verbs ending in /-ɪŋ/. They fall into four groups:

bring, past *brought* /bɹɔt/, unique in this subset of verbs;

ring, *sing*, *spring*, with pasts in /-æŋ/ (*rang*, *sang*, *sprang*);

cling, *fling*, *sling*, *sting*, *string*, *swing*, *wring*, with pasts in /-ʌŋ/ (*clung*, etc.);

a handful of verbs, all uncommon or rare, that belong to the huge

default class. The ones that occur in the senior author's speech are

ding, *king* (in checkers), *ping*, *wing*, pasts *dinged* /dɪŋd/, etc.; at

least some speakers also use *ring* 'put a ring around', past *ringed*, a reviewer notes colloquial British *ming* 'be disgusting', and recently

bling 'be flamboyant' and *bing* 'use Bing' have begun to appear.

The largest group is those with pasts in /-ʌŋ/, and a "commonsense"

approach to morphological change would suggest that that group should

gradually acquire new members. In fact that was long the usual view: "A

form which is statistically predominant is also likely to be productive for

new combinations" (Nida 1949: 45). However, new past tenses in /-ʌŋ/

are conspicuous mostly by their absence in contemporary standard

English, and Yang's hypothesis offers a reason why the situation is stable.

In the senior author's speech N is 15, and the natural logarithm of 15 is

2.70805. The quotient of those numbers is 5.539, and the TP states that no

rule of past tense formation for this group of verbs should become productive if the number of exceptions to it exceeds five. In fact, no matter which rule we choose, the number of exceptions is substantially greater than five: there are eight exceptions to the *ung*-rule, eleven to the default past tense rule, and of course even more to the *ang*-rule. Thus it is not surprising that there is no change in progress. That remains true even for speakers whose only “regular” verb in this group is *wing* (the least rare of those with default past tenses). For them N is 12, and its natural logarithm is 2.48491; the quotient is 4.829, and there are still five exceptions (*bring, ring, sing, spring, wing*) to the rule yielding past tenses in /-ɪŋ/.

At this point it may be convenient to list the productivity thresholds for lexical classes of sizes 3 through 17:

N	θ_N	N	θ_N	N	θ_N
3	2	8	3	13	5
4	2	9	4	14	5
5	3	10	4	15	5
6	3	11	4	16	5
7	3	12	4	17	6

Note that when N is 12 the number of exceptions to the *ung*-rule is very close to the TP threshold. What if a native learner acquires no verbs in /-ɪŋ/ with default past tenses? For such a learner N is 11, but there are

only four exceptions to the rule yielding past tenses in /-ʌŋ/ (*bring, ring, sing, spring*); thus the rule ought to be productive if there are no verbs in /-ɪŋ/ with default past tenses at all. In short, the “*ung*-rule” is a borderline case which might provide an opportunity to test the TP.

In fact one innovative past in /-ʌŋ/ has appeared recently in colloquial English, exemplified in the phrase *they sprung him out of prison* (Margaret Laing, p.c. 27 August 2018). The phrase seems to be U.S. slang, not much more than a century old (see the *Oxford English Dictionary (OED)* s.v. *spring*, v.¹ II.26). We cannot demonstrate that that finite past arose in the speech of a learner who used no verbs in /-ɪŋ/ with default past tenses, but the distribution of facts makes it likely.

Other innovative past tenses that have not caught on can also be explained by the TP. Yang notes that “the productivity of a rule may change during the course of language acquisition” (Yang 2016: 70), depending on the vocabulary that the native learner has learned at any given point. To judge from the 6-million-word corpus of child-directed English in the *CHILDES* database (MacWhinney 2000), the commonest verbs in this class, in descending order, are *bring* (past *brought*), *sing* (*sang*), and *ring* (*rang*), and it is reasonable to suppose that children learn those verbs first more often than not. The seven verbs with pasts in /-ʌŋ/ and *spring* (*sprang*) are subsequently learned in an order that must vary from child to child; verbs with default pasts such as *wing* (*winged*) must

be learned comparatively late, given their rarity. The TP predicts that the productivity of the lexical classes should change as follows. When the only verbs in *-ing* which the child knows are

bring, brought; sing, sang; ring, rang,

the lone exception to the *ang*-rule, *brought*, does not exceed the threshold for $N = 3$, so we expect that rule to be productive, with at least a few children producing past *brang* early in NLA. In fact *brang* does occasionally appear in the acquisition of standard English. If the fourth verb learned is *spring, sprang*, the situation will not change (see the table of thresholds above); if instead the fourth verb is one with a past tense in *-ung*—say, *swing, swung*—the class in *-ang* will remain productive; it will remain productive if *swing* and *spring* are learned in that order as the fourth and fifth verbs in *-ing*, and even if another verb with a past in *-ung* is acquired sixth, so that the verbs in this class are (for instance)

*bring, brought; sing, sang; ring, rang; spring, sprang; swing, swung;
sting, stung,*

given that when $N = 6$ a rule can tolerate three exceptions. However, if the fourth, fifth, and sixth verbs learned all have pasts in *-ung* (and *spring* is acquired only later), it will be the *ung*-rule that is productive. In that case it will make a difference whether the seventh verb acquired is *spring* or another verb with a past in *-ung*; in the former case the *ang*-rule will again become productive, but in the latter case the *ung*-rule will continue

to be productive, since when $N = 7$ a rule can tolerate only three exceptions. When the eighth verb is acquired only the *ung*-rule can be productive, unless and until at least one verb in *-ing* with a default past is acquired. Thus the TP offers an easy explanation for the fact that innovative pasts tenses like *brung* and *sung* are produced less rarely than *brang* in the acquisition of standard English.³

It may be helpful to put this discussion in a broader context. Children learning English virtually always produce default past tenses of “irregular” verbs (see e.g. Pinker 1999: 189–210) but seldom produce innovative non-default past tenses (Xu and Pinker 1995); examples of the latter are rare. The point here is that most of those rare exceptions obey the *ung*-rule.⁴ For instance, in the entire collection of North American child English transcribed in the *CHILDES* database, the past tense of *bring* is correct *brought* 95 times, default *bringed* 6 times, and *brang* or *brung* 9 times. In line with the thought experiment of the preceding paragraph, the non-default but incorrect pasts would not be possible for a speaker with a vocabulary large enough to include most of the verbs in *-ing*, but they are

³ The fact that the past tenses of *sing*, *ring*, *spring* are *sung*, *rung*, *sprung* in some dialects of English comparatively remote from the standard can be explained in more than one way; see fn. 11 below.

⁴ See Yang 2016, ch. 2, for a crosslinguistic review of these and similar findings.

predicted to be spontaneously available to children who have a much smaller vocabulary. In short, we expect *brang* or *brung* to be produced at an early stage of NLA when a learner has not learned enough vocabulary to destroy the transient productivity of the rule. If it is asked why *bringed* is produced comparatively often, given that verbs in *-ing* with default pasts are all rare, we remind the reader of two further facts. In the first place, different learners can be expected to hypothesize different lexical classes; a class of verbs in *-ing* is a universally available hypothesis for children learning English, but it need not be chosen by every learner. Secondly, it is well known that when native learners of English discover the default formation of the past tense they at first overgeneralize it, rejecting the non-default inflections that they have begun to learn, only to return to them a bit later (Marcus et al. 1992). That errors of any kind are rare is a straightforward result of the well-known fact that children are astonishingly good at NLA.

Of course demonstrating in detail that this is the process leading to the production of innovative past tenses for these verbs will involve further and more intensive research on NLA. However, there is another dataset with which the TP can be confronted, namely observed historical changes in a language. Since the history of English over the past six or seven centuries is especially well documented and intensively studied, we believe that it is worthwhile to see whether the TP can help to explain the

English historical data. That is the question this paper will address.

Before we proceed, it seems advisable to clarify a methodological point. We noted above that different native learners can be expected to hypothesize different lexical classes, though some will be hypothesized more often than others; only experimental work on NLA can shed further light on that point. Because we are dealing with the distant past, we are constrained to adopt a different strategy: we start from changes that have in fact occurred in the history of English and attempt to retrofit the analysis so as to explain those developments in accordance with the TP. This approach is complementary to simple computational models of inductive learning, such as those constructed by Yip and Sussman 1997 and Albright and Hayes 2003, which operate on the phonological representations of words to produce rules such as the *-ung* rule. In either case some further hypothesis, such as the TP, is needed to assess the productivity of rules.

If we could explain everything that we observe by our strategy, we could say nothing about the limits of the TP. However, it will be seen that there are observed developments which the TP probably cannot explain; most importantly, they are precisely the developments which do not seem to be rule-governed and might have been initiated by adults.

3. The 16th and 17th centuries.

Over the course of the 16th and 17th centuries written English becomes increasingly uniform, but there is still a good deal of variation. In the context of that variation several noteworthy innovations in past tense formation occurred. *Stringed* was replaced by *strung*, suggesting that the *-ung* rule was productive for some speakers for at least part of that period; more surprisingly, *digged* was replaced by *dug* and *sticked* was replaced by *stuck*, suggesting that the rule was actually broader, taking as input verbs with /ɪ/ in the root followed by any single velar consonant (or homorganic velar cluster—see further below). We need to determine whether such a rule could have been productive according to the TP, or whether some other type of change must have occurred.

The two principal resources for investigating the development of English during this period are the *OED* and the *Penn-Helsinki Parsed Corpus of Early Modern English* (*PPCEME*; Kroch et al. 2004). The *OED* records every form attested and attempts to give the earliest instance of each attested form, usually with (at least approximate) success. The *PPCEME*, though it includes more than 1.7 million words of running text, is not large enough to offer examples of every lexeme, let alone every form; on the other hand, it gives a fair indication of which lexemes occur most frequently, and for common lexemes it provides good information about the relative frequency of competing forms.

We also consulted the *Parsed Corpus of Early English Correspondence* (*PCEEC*; Nevalainen et al. 2006), but it contained too few relevant forms to contribute significantly to the relevant statistics; for instance, though *sing* is one of the commonest relevant verbs in the *PPCEME*, its finite past is unattested in the *PCEEC*, which also contains only two examples of its past participle *sung*. Forms from the *PCEEC* will be cited in the footnotes where relevant. Other potential corpora will also be mentioned where relevant.

4. New past tense forms in Early Modern English.

The currently ascertainable facts regarding the new past tense forms noted in the last section are the following.

The *OED* records a single instance of the verb *string* from late Middle English (ME), *with bowes gode wel y-strenged* ‘with good bows well strung’, from the *Laud Troy Book*, ca. 1425; the verb *to string* (a bow or a musical instrument) next appears in the 16th century, the earliest *OED* quotations being from 1530 and 1545.⁵ Since the verb is clearly derived

⁵ There are also a couple of quotations each in the more obscure meanings ‘to furnish (a garment) with ties’ (first in 1548) and ‘to remove the notochord from (a lamprey)’ (1508). It is not clear whether those meanings are relevant; in any case they are rare, and the *OED* quotes no

from the noun *string* (which was inherited from Old English (OE), first attested in the 8th- or 9th-c. poem *Andreas*), one would expect it to inflect according to the default pattern, with a past and past participle *stringed*. In fact *stringed* and *strung* are both quoted from the 16th c. onward, though *stringed* is gradually eliminated, falling out of use in the 18th c. In the 16th c. *stringed* is first quoted from 1530, and again from 1548 (past ptc. in both instances); *strung* (likewise past ptc.) is first quoted from Spenser (*Virgil's Gnat*, 1591), and from Donne (1612), Chapman (1613), Shakespeare (before 1616), and obscure authors of the 1590's and 1600's. The past and past participle of this verb are not attested in the *PPCEME*. Though the attestation is too sparse for absolute certainty, it appears that *stringed* was in place early in the 16th c. and began to be replaced by *strung* in literary English sometime in the second half of the century.

For *dig* the record is a bit fuller and the situation slightly clearer. The verb was borrowed from French around 1300, and a default past tense *diggede* and past participle *digged* are well attested in late ME. Such forms continue to occur throughout the 16th and 17th centuries; the *PPCEME* has examples from Tyndale's bible (1520's and '30's, adopted without change in the King James translation of 1611), from the works of Sir Francis Bacon, William Clowes, and Richard Hooker (all late 16th c.),

past tense for either. The now common meanings 'to put (things) serially on a string' and 'to arrange in a line' both appear first in the 17th c.

from the letters of Lady Katherine Paston, John Taylor, and Henry Oxinden (all latest 16th and early 17th c.), and from statute books throughout the period. Early examples of *dug* happen not to occur in the corpus, but the *OED* (under *dig*, section II.4) gives an example of past ptc. *dug* from 1580, and another occurs in the *PPCEME* sample from Pepys' diary (1660's). The finite past *dug* does not seem to be attested before the beginning of the 18th c.; the *PPCEME* records *dug* in both functions in the memoirs of Celia Fiennes, completed ca. 1702. It appears that the past ptc. *dug* was created around the same time as *strung* but either spread more slowly or was resisted in written English; so far as our evidence goes, the finite past *dug* could have been created in the 17th century.⁶ Both had become acceptable among the literate class by the end of that century.

For the past of *stick* we have much more information. It is the descendant of OE *stician* 'to pierce; to adhere', whose past tense *sticode* would be expected to develop into a default past tense *sticked*. Sure enough, a past *sti(c)ked(e)* is well attested in ME and continues to be attested in the 16th century. The *PPCEME* includes two examples, both finite pasts: *stycked* from a letter of Sir Thomas More,⁷ who was born in

⁶ This is a principal reason for using 17th-c. data from the *PPCEME*, even though *dug* makes its first appearance late in the 16th.

⁷ Also in the *PCEEC*.

1478, and *sticked* from the works of William Clowes, who was born in the early 1540's. But *stuck* is attested significantly earlier than *strung* or *dug*. The earliest attestation in the *PPCEME* is from *Ralph Roister Doister*, a play written by Nicholas Udall (born 1504) in about 1552; it is a finite past *stucke*, rhyming with *Gawyn Goodlucke* (a major figure in the play). Another seven examples, both finite past and past participle, from later in the century occur in the corpus.⁸ It seems clear that by the early 17th century the innovative strong forms were normal.

In the case of *stick* there is a further fact which complicates the picture. A synonymous strong verb *steke(n)* also appears in ME; it is not attested in OE,⁹ yet it appears to be cognate with Old Frisian *steka*, Old Saxon *stekan*, and Old High German *stehhan*, all meaning 'pierce'. In late ME the past tense of this verb was *stake*, well attested in the works of Malory (also spelled *stak* and *stack*; the reconstructable pronunciation is /sta:k/), and the *OED* records a past participle *stoken* ~ *ystoke*; the inflection of the verb was precisely parallel to that of *break* (*brake*, *broken*). Just as the

⁸ The *PCEEC* examples are all from the 17th century.

⁹ That is probably because it had been lost in all the OE dialects for which we have adequate attestation but survived in one or more dialects that are unattested (or nearly so). For various historical reasons the dialects of almost the whole Midlands area are very sparsely attested or completely unattested in the OE period.

vowel of the latter verb's past participle was levelled into its finite past, yielding *broke*, one would expect to find a past *stoke*, and the *OED* records such a form from the 15th and 16th centuries. However, NONE OF THESE FORMS is likely to be the direct ancestor of *stuck*. If a form survives intact from one generation to the next, the only kind of change it undergoes is "sound change", i.e. spontaneous changes in pronunciation. The historical record shows that sound change is overwhelmingly regular (and modern sociolinguists have identified a process of sound change that is completely regular; see e.g. Labov 1994 and forthcoming, Fruehwald 2013). Apparently irregular changes in pronunciation which we cannot explain do occasionally appear in the record (though much more rarely than nonspecialists suppose), but to posit such a change without overwhelming evidence is always to flout the odds.¹⁰ No professional gambler would do that, and no historical linguist should. Since none of the vowels in the various forms of *steke* regularly became /ʌ/, we must conclude that none of them can be the ancestor of *stuck*. Whether the existence of a synonymous inherited strong verb is relevant to the creation of *stuck* in some other way will be considered below.

¹⁰ This is why we discount the suggestion of Dobson 1968: 509, 676 that *stroke* could be the ancestor of *struck* (on which see further below); of course that is not completely impossible, but the odds are against it.

5. The productivity of /ʌ/ in Early Modern English: the case of *strung*.

Making sense of the changes under consideration involves dealing with several relevant questions more or less simultaneously. We present the discussion in an order which we hope will make the situation maximally intelligible.

First of all, we need to address the complicating issue of past participles, because the situation among the past participles is not the same as the situation among the finite pasts in standard Modern English. Since finite pasts in *-ang* (*sang, rang, sprang*) correspond to past participles in *-ung* (*rung, sung, sprung*), there are actually 10 past participles in *-ung* in standard Modern English—so even for a speaker who has 15 verbs in *-ing*, we might expect the *ung*-rule to be productive for past participles, since the TP predicts that the *ung*-rule should be able to tolerate five exceptions (in the senior author's case, *brought, winged, dinged, kinged, and pinged*). In nonstandard English the *ung*-rule probably is marginally productive for past participles, since *brung* does occur, but in standard English it has not replaced *brought*, and past participles like “*wung*” (for *winged*) are not attested anywhere (so far as we know). Since the TP otherwise accounts well for the productivity of morphological rules, there must be a specific reason for its inapplicability in this case. We might suggest that the paradigmatic relationship between the finite past and the past participle is

responsible; in effect, since finite past *winged* is not under pressure from “*wung*”, neither is past participle *winged*. The great frequency of *brought* might also be a factor; so might the default status of *winged*, etc. More research on this problem is needed, but we do not see that it jeopardizes our investigation into the finite pasts under discussion.

Secondly, while the past participles of verbs in *-ing* in the 16th c. were essentially identical to their modern descendants (with the exception of *stringed*), the finite past tenses were not. The finite pasts *sang*, *sprang*, *rang* were in competition with *sung*, *sprung*, *rung*, and that competition was vigorous.¹¹ The *PPCEME* includes twelve examples of *sang(e)* and twelve of finite past *sung* from the beginning of the 18th c. or earlier; most

¹¹ This competition arose in the Middle English period. In Old English these verbs had two finite past stems; for instance, the indicative 1sg. and 3sg. of *singan* were *sang*, but the 2sg. was *sunge*, the plural was *sungon*, and the subjunctives were sg. *sunge*, pl. *sungen*. In the north *sang* became the only finite past in the 13th c., but further south both stems survived for another century or more, and different dialects levelled them in one direction or the other (see e.g. Brunner 1948: 76–7). London English acquired both finite pasts because of the massive influx of migrants from other areas of England leading to substantial dialect mixture (Ekwall 1956). For further discussion see especially Lass 1994, Cheshire 1994, both with references.

remarkably, their excerpts from the diary of Samuel Pepys—written by a single individual over the course of about a decade—include two of each. Though the finite past tense of *spring* is not as common, the situation is comparable: there are six examples of *sprang(e)* (ignoring a duplicate in the King James bible) and four of *sprung*. The finite past *rang* does not occur in that corpus, but *rung* occurs twice. It seems clear that for at least some speakers no fewer than ten verbs had both finite pasts and past participles in *-ung*; the only exceptions were *bring*, *string*, and a handful of very rare verbs like *wing* ‘to dismember (a partridge)’. For those learners who encountered only one or two verbs in *-ing* with default pasts other than *string*, N was 12 or 13 and there were only three or four exceptions to the *-ung* rule (*bring*, *string*, and the other default-past verb(s)); in their acquisition of English the replacement of *stringed* by *strung* is exactly what the TP would lead us to expect, since a rule with twelve inputs will tolerate four exceptions and a rule with thirteen inputs will tolerate five. Even learners who heard both *sang* and *sung*, etc., as finite pasts in the speech of some adults (like Samuel Pepys) MIGHT have been moved to replace *stringed* by *strung* if the TP is not inhibited by competing forms (a question which we are not yet in a position to answer definitively, though see Sneller, Fruehwald, and Yang 2019 for suggestions).

In short, the TP actually actually leads us to expect replacement of

stringed by *strung* in the 16th c. (if not before). Presumably *brought* escaped replacement by *brung* because its extreme frequency made it easier to learn and more resistant to change; rarer verbs with pasts in *-inged* were perhaps not learned during NLA, but only later in life.

However, when *sang*, *sprang*, *rang* won their respective competitions in standard English—toward the end of the 17th c., to judge from the available evidence—the *ung*-rule could no longer be productive in that dialect, because the number of finite pasts in *-ung* dropped to only 7, while the number of input verbs remained the same (or perhaps increased); as the table of thresholds indicates, a rule with 12 inputs should not tolerate 5 exceptions ($12 - 7$), nor should a rule with 13 inputs tolerate 6 exceptions ($13 - 7$). The *ung*-rule might have remained productive in other dialects, and occasional speakers continue to produce forms in *-ung* even now, but there was no longer any chance that such forms would “catch on” in the standard language.

Innovative *strung*, then, is fully consistent with the TP. *Stuck* and *dug* are more complicated cases.

6. *Stuck*.

While there were many verbs with finite past tenses in *-ung* in London English in the years before 1500, there were apparently none with finite past tenses in *-uck*. Therefore if *stuck* arose by productive rule, the rule

must have applied to more than verbs with roots in *-ick*. The most conservative alternative, requiring the fewest additional hypotheses, would seem to be verbs with roots in which /ɪ/ was followed by a single velar consonant, but there is a complication. ModE word-final /-ŋ/ reflects early ME /-ng/ [-ŋg], and we need to consider when word-final -g after a nasal was lost. Unfortunately the loss seems to have begun in eastern Midlands dialects in the 14th century, but not to have been complete in educated English until the late 16th (Dobson 1968: 963–4).¹² Since we cannot pinpoint exactly when and where *stuck* was created, and since it could have spread into London English from a dialect in which the phonotactics were different, we must consider whether the TP can account for its creation under two different scenarios, one based on applying the *ung*-rule

¹² We are grateful to an anonymous reviewer for pointing this out. — Dobson’s summary of the facts is indispensable, but some details of his discussion need revision. To speak of a “standard” pronunciation of English even now is to idealize a good deal, and for the 16th century any such idealization seems misconceived (Milroy 1994). Nor is the fact that the early orthoepists do not recognize the velar nasal as a separate sound decisive; Dobson recognizes that their interpretations of spellings are not entirely trustworthy in his discussion of *-mb* (p. 960), and the same reserve seems all the more appropriate in a discussion of *-ng*, given that that was the only conceivable way of spelling /ŋ/.

to verbs with roots in /ɪ/ plus any single velar (for dialects in which /-ŋ/ already existed) and one based on applying the rule to verbs with roots in /ɪ/ plus any single velar or homorganic velar cluster (i.e. /-ŋg/ or /-nk/).

We present the options in that order.

Figuring out which verbs in /-ɪK/ were present in the speech heard by a typical native learner in, say, 1510 is probably beyond the capability of philology; we will need to employ reasonable proxies. One plausible proxy is the collection of verbs whose finite past tenses and past participles are actually attested in the 16th- and 17th-c. samples in the *PPCEME*.¹³ Here is the list, with numbers of attestations:

brought (scores of exx., sometimes more than 10 in a single sample);
digged (11x), *picked* (7x), *pricked* (6x), *thicked* (4x, meaning
‘thickened’), *sticked* (2x), *nicked* (1x), *rigged* (1x);
sang/sung (36x), *sprang/sprung* (12x), *flung* (10x), *stung* (5x) / *stinged*

¹³ A reviewer points out that one might also construct such a list from a concordance to Shakespeare’s plays, which are largely in dialogue, or the *Corpus of English dialogues 1560–1760* (Kytö et al. 2006). Both should be attempted, but it is not clear that either would be a better fit to child-directed speech, which tends to include a language’s most frequent lexemes overall—making the sheer size of a corpus the most important variable.

(1x),¹⁴ *rung* (3x), *slung* (2x), *wrung* (2x)

—plus *stuck* (8x) and *dug* (3x), not present at the beginning of the period.

In addition (as a reviewer reminds us), there is *strick*, a rare variant of *strike*, which for at least one author who used it had a past *strake* but for others had a past *struck* (see section 7 below).

Thus for this calculation there are 15 or 16 verbs altogether, and only 7 or 8 follow the rule that was extended to yield *stuck* and *dug*, so that there are 8 or 9 exceptions. Since a rule potentially applying to 15 or 16 lexemes should tolerate only 5 exceptions, those 8 or 9 should easily be enough to prevent it from becoming productive. However, it is plausible to suppose that not all these forms were acquired in NLA. In particular, *thicked* occurs only in discussions of the preparation of cloth (in statutes of the period), *rigged* describes the outfitting of a ship, and a reviewer observes that *nick* is also quite rare in 16th- and 17th-c. documents. If we remove those verbs, the total is 12 or 13; a class of 13 should tolerate only 5 exceptions, and a class of 12 only 4. In the latter case we have 5 exceptions, namely *brought*, *digged*, *picked*, *pricked*, and *sticked*; in the former (including *strick*) we have 6 exceptions (including *strake*) or 5 (if

¹⁴ *Stinged* must be a nonce form, an outlying dialect form, or a personal idiosyncrasy; the irregular *stung* is clearly usual, and had been for some generations.

the past was *struck*). Only under the last scenario does the TP predict the creation of *stuck*. This is a borderline result: only if *struck* was already in existence does the TP predict *stuck*. We do not really know whether that was the case or not, but the chronology of attestations suggests that *stuck* was created first, in which case the TP fails in this case too.

Of course we might be using the wrong proxy. An alternative proxy would be a relevant modern list of lexemes in child-directed speech, adjusted so as to account for the differences in inflection between early 16th-c. English and present-day English. In the 6-million-word *CHILDES* corpus of child-directed American English a dozen relevant verbs appear in the past tense at least once per million tokens each; we give them here with their early-16th-c. past forms:¹⁵

1	bring, brought	7	fling, flung
2	sing, sang/sung	8	kick, kicked
3	ring, rang/rung	9	lick, licked
4	spring, sprang/sprung	10	pick, picked
5	swing, swung	11	stick, sticked
6	sting, stung	12	dig, digged

Half of these dozen verbs had past tenses in *-ung* in the early 16th century,

¹⁵ These and other classes of English verbs in the *CHILDES* corpus were collected and extensively used to account for English NLA in Yang 2016; see Yang 2016 *passim*.

but according to the TP the six verbs in this list that have past tenses of other kinds should have been enough to prevent the formation of *stuck* in NLA, or at least to prevent it from catching on, if these verbs are a representative sample of child-directed speech in the 16th century.

But what if *stuck* arose in a dialect in which *bring*, etc. were still [briŋg], etc.? In that case verbs in *-ink* need to be included in the potential lexical class membership. Here is the list of past tenses attested in the *PPCEME* with verbs in *-ink* included (but *stinged* omitted, see above):

brought, thought (both very common);

digged (11x), *picked* (7x), *pricked* (6x), *linked* (5x), *thicked* (4x),

sticked (2x), *nicked* (1x), *rigged* (1x), *winked* (1x), *blinkt* (1x);

drank/drunk (44x),¹⁶ *sang/sung* (36x), *sprang/sprung* (12x), *flung*

(10x), *stung* (5x), *rung* (3x), *shrank/shrunk* (3x), *slung* (2x), *wrung*

(2x), *stank/stunk* (2x), *sank/*sunk* (2x)¹⁷

—plus *stuck* (8x) and *dug* (3x), not present at the beginning of the period, and possibly *strake* (see above).

¹⁶ Though both these forms occur as finite pasts, *drank* appears 41 times and *drunk* only 3; thus the balance is very different than for other verbs with such forms in competition.

¹⁷ *Sunk* happens not to occur as a finite past in the *PPCEME*, though *OED* citations show that such a form was in use. (*Sunk* does occur as a past participle 14x in the *PPCEME*.)

We now have a class of 23 or 24 verbs, so any rule should tolerate 7 exceptions; but there are at least 11 exceptions to any potential rule. If we exclude *thicked*, *rigged*, *nicked* (see above) and *blinkt*—the single token of the last means ‘turned sour’, not likely to occur in child-directed speech—we still have 19 or 20 verbs, so any rule should tolerate 6 exceptions, but there are still at least 8. Once again the TP does not predict *stuck*.

Finally, let us repeat the experiment using relevant verbs from the CHILDES database that are not rare, this time including verbs in *-ink*, with their 16th-c. past tenses, namely:

- | | |
|-------------------------|-------------------|
| 1 bring, brought | 9 sting, stung |
| 2 think, thought | 10 fling, flung |
| 3 sing, sang/sung | 11 kick, kicked |
| 4 ring, rang/rung | 12 lick, licked |
| 5 spring, sprang/sprung | 13 pick, picked |
| 6 drink, drank/drunk | 14 stick, sticked |
| 7 sink, sank/sunk | 15 dig, digged |
| 8 swing, swung | 16 blink, blinked |

The last of these verbs might not have been nearly so common in the 16th century, but *wink*, *winked* was probably commoner, since it could mean ‘keep the eyes closed’; the numbers should be approximately correct in any case. We now have 16 verbs, among which any rule should tolerate 5 exceptions, but the *ung*-rule applies to only half of them —there are still 8

exceptions.

In short, we have not certainly been able to derive *stuck* via the application of the TP in NLA under any reasonable assumptions about what 16th-c. learners of English were hearing (though there is a bare possibility that we might be able to). It seems at least as likely that *stuck* was created by some other process, and in fact we can propose one. As we noted at the end of section 4, there was beside the weak verb *stick* (past *sticked*, ptc. *sticked*) a synonymous strong verb *steke* (*stake*, *stoken*; later *stoke*, *stoken*). It seems reasonable to suggest that some speakers—not necessarily native learners—confused the two and were therefore prompted to create a strong past and past participle for *stick*. But given that the vowel in the root of the later was /ɪ/ (throughout the word's attested history), a strong past or past participle could only be either *stack* /stak/ (which would develop into /stæk/ by regular sound change) or *stuck* /stʊk/ (which would develop into /stʌk/). In the event *stuck* was created in both functions and caught on. This could be called lexical analogy, though it is lexical analogy of a special kind: not the influence of a lexeme on a quite different lexeme that happens to resemble it in various ways, but the perhaps mutual influence of very similar (and in fact etymologically related) lexemes competing for position as the default

representation of a meaning in the basic vocabulary.¹⁸ Whether such a process can be modelled as rule-governed is doubtful; we have not been able to come up with a rule which the TP could test.

There is another piece of evidence which points in the same direction. We have had to deal with it in the calculations in this section, but we have not discussed it in detail because it adds a further layer of complexity to the picture. We must do so before we turn to *dug*.

7. *Strike / strick, past stroke / strake / struck.*

The OE verb *strīcan* ‘to stroke, to wipe; to dash, to run’ is the direct ancestor of Modern English *strike* (the meaning has evolved considerably over time). Though only the present is attested in OE, the past 3sg. can only have been **strāc* and the past participle **stricen*, parallel to *writan*, *wrāt*, *writen*, because the inflection of OE strong verbs with *ī* in the root is

¹⁸ A more typical example of lexical analogy from the history of English verb inflection is the early ME creation of past *cauzte* (beside *cacchede*) to the French loanword *cacchen* on the model of *lauzte*, the past of the synonymous native verb *lacchen* (see the *OED* s.vv. *catch*, v. and *latch*, v.1). In principle such straightforward lexical analogies can be modelled as rules. Of course a rule covering only two lexemes is comparatively inefficient in an adult lexicon, but in the restricted lexicon of a native learner that need not be the case.

uniform. The result should have been *strike*, *stroke*, *stricken* in the 16th c., parallel to *write*, *wrote*, *written*, and in fact that paradigm is well attested. The present stem *strike* is common; perusal of about 40% of the *PPCEME* yielded 21 examples. Past *stroke* appears seven times in the same corpus, and the examples are distributed throughout the 16th c. and down to 1630. Past participle *stricken* occurs with some frequency (and survives in Modern English, though more often as an adjective than as a participle).¹⁹

But there are other forms in competition with all of the above in the *PPCEME*. There is a rare present *strick* beside *strike*. It too was a descendant of OE *strīcan*, but one in which the vowel had been shortened; that is known to have happened to a handful of class I strong verbs in some dialects of OE (Seebold 1966). One example of *strick* in the *PPCEME* occurs in the work of Thomas Harman, published in the 1560's:

“A vaunt verlet,” quoth this vpright man, and letes dryue with all his force at this hosteler, and after halfe a dosen blowes, he *strycks* his staffe out of his hande,

(*PPCEME*, HARMAN-E1-P1, 64.7–9)

Much more surprisingly, there are two examples from John Locke's well-known treatise on education, published in 1693. Since Locke was born in

¹⁹ Past participle *stroke* also occurs, e.g. in one of the Cely letters from the late 15th century and in a 17th-century letter of John Chamberlain, both in the *PCEEC*.

Somerset in 1632, it is possible that *strick* is a dialect form that he never abandoned, but apparently it was not stigmatized. So far as our evidence goes, *strick* persisted as a variant of *strike* in some areas, or in the speech of some individuals, into the 18th c.²⁰

It seems clear that *struck* must have been created as a past tense to *strick* in much the same way that *stuck* was formed to *stick*—and since *strike* / *strick* was already a strong verb, *struck* presumably replaced *stroke* and *stricken*. But already by the 16th c. the pattern of variation was more complex. The passages from Locke's works in the *PPCEME* contain no example of a past tense or past participle of *strick*, but an online search of Locke's works on Project Gutenberg turns up examples of *struck* in both functions, as one would expect. However, for Thomas Harman the past of *strick* was *strake*; the passage quoted above continues:

and as this hosteler stept backe to haue taken vp his staffe agayne, his glymmeringe Morte flinges a great stone at him, and *strake* him one the heade that downe hee fales, wyth the bloud about his eares,
(*PPCEME*, HARMAN-E1-P1, 64.200–1)

The past *strake* also occurs in two *PPCEME* passages from the chronicle of Robert Fabyan, who died in 1516, beside a past participle *stryken* (i.e., *stricken*); his present for this verb is not attested in the *PPCEME*, but an

²⁰ The *PCEEC* seems to have only one example, from a letter of Nicholas Bacon in the second half of the 16th century.

online search of the Open Library copy yields four examples of *stryke* (as well as many more of past ptc. *stryken*). The *PPCEME* attests both present *strike* and past *strake* from the biography of Sir John Perrott, written in the 1590's. But many authors whose present stem is *strike* already have a past *struck*, exactly as in Modern English. The diary of Henry Machyn, written in the 1550's and '60's, exhibits a paradigm *stryke, struck, stryken*. The examples of *struck* next in chronological order are from the chronicle of John Hayward, written late in the 16th c. or early in the 17th, and from a comedy by Thomas Middleton published in 1630; both are past participles. No finite past other than *struck* appears in the texts of the *PPCEME* from after 1630, though the past participles *struck* and *stricken* continue to compete in general use until at least the 1660's.²¹ Samuel Pepys and Celia Fiennes both have the modern paradigm *strike, struck, struck*.

The reader who finds this picture confusing is no more clueless than we are; it seems clear enough that by the middle of the 16th c. the competition between stems of *strike / strick* could be described as a free-for-all. That is relevant because influence of *stick* on *strike / strick* and

²¹ Forman's diary, which covers the second half the 16th c., yields a variant *stroken*. Past *strake* and past participle *strucken* occur several times in a wide range of letters in the *PCEEC*, adding to the picture of extensive variation in several dimensions.

vice versa is a real possibility, given the similarity of form and meaning that the two verbs exhibit: in effect, the entirely new strong past *stuck* and the remodelled strong past *struck* might have supported one another. The fact that *strike* seems to have replaced *strick* while on the other hand *struck* replaced *stroke* makes the entire situation difficult to judge. In any case there seems to be no rule that the TP could be used to assess, and once again we cannot be certain that these innovations occurred in NLA at all; it is possible that adult native speakers were responsible for them.

8. *Dug*.

The pattern of evidence suggests, though it does not quite establish, that *dug* was created later than either *struck* or *stuck*; it might or might not have been created later than *strung*. Let us therefore revisit the first calculation under section 6, with *thicked*, *rigged*, and *stinged* (see fn. 14) excluded but *struck* and *stuck* included. We have the following set of relevant forms for native learners who hear *strick* in use:

brought;

digged, picked, pricked, nicked;

sang/sung, sprang/sprung, flung, struck, stuck, stung, rung, slung,

wrung.

In this case *N* is 14, which should tolerate five exceptions; since we have only five exceptions to the rule “replace /ɪ/ with /ʊ/”, we expect the rule to

be productive. The only surprise is that *digged*, the commonest default-past verb of this set in our sample, has been targeted for replacement; that will have to be laid to contingent factors no longer recoverable. For native learners hearing *strike* in use, we will have to omit *struck* from the above list, and it is possible that *nicked* should be omitted even early in the 17th century. In either case N is then 13, which also tolerates five exceptions, or, if both are omitted, N is 12, which tolerates four, so the rule should have been productive even for those learners.

In short, if *dug* was the last of these innovative forms to be constructed, the TP predicts its appearance. Moreover, just as the victory of *sang*, *sprang*, *rang* late in the 17th c. explains why *winged* was not replaced by “*wung*”, it also explains why *rigged* did not yield to “*rug*”—nor *picked* to “*puck*”, and so on, in accordance with the TP. Note that the verbs in *-ick* which entered English after the victory of *-ang* over *-ung* have all adopted and retained the default inflectional pattern; to judge from the *OED* entries, they include at least *click*, *flick*, *tick*, and *trick*. Finally, the fact that the calculations in this section worked out so neatly suggests that a list of forms most frequently attested in non-specialized texts is a reasonable proxy for the unobtainable list of input forms for NLA in centuries long past; for relevant discussion see now Kodner 2019, 2020. Of course this conclusion regarding the origin of *dug* cannot strictly be proved, but it is a construct *où tout se tient*.

9. Looking forward.

The results just described are clear-cut enough and interesting enough to prompt testing of the TP against other instances of past tense “irregularization” in English. While a full study of any of the relevant verbs is beyond the scope of this paper, two cases seem clear enough that suggestions might be useful.

The replacement of *dived* by *dove* in 19th-c. North American English appears to be easily explainable by the TP. So far as we can discover, only four verbs in *-ive* /-aɪv/ were common enough to be learned regularly by small children at that time and place;²² their finite pasts were:

drove, strove;

throve ~ thrived;

dived.

In addition, there were two verbs of more specialized meaning, whose finite pasts were:

hived (bees), rived (cedar shakes with a froe).

Even if small children learned the specialized verbs, pasts in *-ove* could have become productive in the NLA of those learners who heard *throve*, according to the table in section 3; for children who did not learn the

²² *Shrive* and *swive* were no longer in everyday use; *jive* had not yet been coined.

specialized verbs early, the *-ove* rule could have become productive even if they heard *thrived*. Thus the TP easily predicts the appearance of *dove*.

A contrary case is *hung*. Like *strike, struck* (see section 7), *hang, hung* (transitive) has been a strong verb throughout its history, but its finite past is not the one that is expected etymologically. To judge from the citations in the *OED*, *hung* might first have appeared in the 16th c.; on the other hand, some of the Middle English finite pasts spelled *hong(-)* might conceivably have contained /ʊ/ in the root. This case requires a thorough study in detail. However, it is already clear that the verb had several competing finite pasts in Middle English, and it would be reasonable to wager that the history of *hang* will turn out to be much like that of *strike*—not involving NLA nor the TP.

There is also a relatively recent phenomenon that seems to be different from both the phenomena dealt with in this paper. A finite past *snuck* (to *sneak*) first appears in the USA in comic writing in the mid-19th c.; William Faulkner, toward the end of his short story “Was”, has a character say *you skun the hen-house one time too many*; Dizzy Dean is quoted as having said *he slud into third*;²³ a collection of oral histories of D-Day includes the statement *those that were wounded we drug up behind* (Astor 1994: 196); and a finite past *rutch* (to *reach*) was heard at a sporting event at Pennsylvania State University (Philip Baldi, p.c.; several of these forms

²³ We are grateful to Patrick Stiles for calling this to our attention.

are also adduced by Cheshire 1994: 122). It seems clear that /ʌ/ is spreading as a finite past tense vowel for verbs in North American English, but it is not clear exactly what is going on, nor why default-pattern verbs occasionally acquire such finite pasts. The discussion of Hogg 1988 and especially Cheshire 1994, with references, is highly relevant (and Cheshire's observations must be largely correct), but more information about the process is needed. That is a topic for future research.

10. Conclusion.

A close examination of the “irregularization” of verbs in Early Modern English confirms the usefulness of the Tolerance Principle in predicting which morphological rules will become productive. It also shows that the TP is not the whole story, as might have been expected. It appears that situations in which synonymous and similar but differently inflected verbs, such as *steke* and *stick*, are in competition can lead to unexpected results; competition between multiple stems of a single verb, as in the case of *strike* / *strick*, can also yield unforeseen anomalies. It is very likely that adult native speakers are responsible for those innovative forms.

References.

Albright, Adam, and Bruce Hayes. 2003. Rules vs. analogy in English

- past tenses: a computational/experimental study. *Cognition* 90(2): 119–61.
- Astor, Gerald. 1994. *June 6, 1944: the voices of D-Day*. New York: St. Martin's.
- Berko, Jean. 1958. The child's learning of English morphology. *Word* 14: 150–77.
- Brunner Karl. 1948. *Abriß der mitttelenglischen Grammatik*. 2nd ed. Halle: Niemeyer.
- Cheshire, Jenny. 1994. Standardization and the English irregular verbs. In Stein and Tieken-Boon van Ostade (edd.) 1994: 115–33.
- Dobson, E. J. 1968. *English pronunciation 1500–1700*. Oxford: Clarendon Press.
- Ekwall, Eilert. 1956. *Studies on the population of medieval London*. Stockholm: Almqvist & Wiksell.
- Fruehwald, Josef. 2013. *Phonological influence on phonetic change*. Philadelphia: U. of Pennsylvania dissertation.
- Hogg, Richard M. 1988. Snuck: the development of irregular preterite forms. In Nixon, Graham, and John Honey (eds.), *An historic tongue: Studies in English linguistics in memory of Barbara Strang* 31–40. London: Routledge.
- Kodner, Jordan. 2019. Estimating child linguistic experience from historical corpora. *Glossa* 4(1): 122. DOI:

<http://doi.org/10.5334/gjgl.926>.

- . 2020. *Language acquisition in the past*. Philadelphia: U. of Pennsylvania dissertation.
- Kroch, Anthony, Beatrice Santorini, and Lauren Delfs. 2004. *The Penn-Helsinki parsed corpus of early Modern English*. Philadelphia: Dept. of Linguistics, U. of Pennsylvania.
- Kytö, Merja, Jonathan Culpepper, et al. 2006. *A corpus of English dialogues 1560–1760*. Uppsala: Dept. of English, Uppsala Universitet.
- Labov, William. 1994. *Principles of linguistic change*. Vol. 1. *Internal factors*. Oxford: Blackwell.
- . Forthcoming. The regularity of regular sound change. To appear in *Language*.
- Lass, Roger. 1994. Proliferation and option-cutting: the strong verb in the fifteenth to eighteenth centuries. In Stein and Tieken-Boon van Ostade (edd.) 1994: 81–113.
- MacWhinney, Brian. 2000. *The CHILDES project: tools for analyzing talk*. 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates.
- Marcus, G. F., et al. 1992. *Overregularization in language acquisition*. Chicago: U. of Chicago Press.
- Milroy, James. 1994. The notion of “standard language” and its applicability to the study of Early Modern English pronunciation. In

- Stein and Tieken-Boon van Ostade (edd.) 1994: 19–29.
- Nevalainen, Terttu, et al. 2006. *The parsed corpus of early English correspondence*. Text version. Helsinki: U. of Helsinki.
- Nida, Eugene. 1949. *Morphology: the descriptive analysis of words*. 2nd ed. Ann Arbor: U. of Michigan Press.
- Pinker, Steven. 1999. *Words and rules: the ingredients of language*. New York: Basic Books.
- Seebold, Elmar. 1966. Die schwundstufigen Präsentien (Aoristpräsentien) der *ei*-Reihe. *Anglia* 84.1–26.
- Sneller, Elizabeth, Josef Fruehwald, and Charles Yang. 2019. Using the Tolerance Principle to predict phonological change. *Language Variation and Change* 31(1): 1–20.
- Stein, Dieter, and Ingrid Tieken-Boon van Ostade (eds.). 1994. *Towards a Standard English 1600–1800*. Berlin: Mouton de Gruyter.
- Strik, Oscar. 2015. *Modelling analogical change: A history of Swedish and Frisian verb inflection*. Groningen: U. of Groningen dissertation.
- Xu, Fei, and Steven Pinker. 1995. Weird past tense forms. *Journal of Child Language* 22(3): 531–56.
- Yang, Charles. 2016. *The price of linguistic productivity: how children learn to break the rules of language*. Cambridge, MA: MIT Press.
- Yip, Kenneth, and Gerald J. Sussman. 1997. Sparse representations for fast, one-shot learning. *Proceedings of the National Conference on*

Artificial Intelligence 521–7.