

Hans-Olav Enger

Type frequency is not the only factor that determines productivity, so the Tolerance Principle is not enough

<https://doi.org/10.1515/bgsl-2022-0013>

Abstract: Inflection classes that have many members often gain members from classes that have fewer. While this tendency is often pointed out in diachronic linguistics, the American psycholinguist Charles Yang (2016) goes further. He claims this to be always the case, so that minority classes cannot be productive at the expense of majority classes, and that productivity actually can be predicted. By this view, productivity is a direct function of type frequency; there are no other factors determining whether a pattern is productive.

The claim of this paper is that type frequency is not the only factor determining productivity, and that while Yang's approach, the ›Tolerance Principle‹, is interesting, it cannot be upheld in its present form. The paper presents an example of suppletion spreading in Germanic, and it presents examples of minority patterns spreading in North Germanic. Parallels outside of Germanic are pointed out. Also, it is argued that Yang's (2016) analysis of English verb inflection and German noun inflection is insufficient, so these important case-studies, presented in favour of the Tolerance Principle, do not support it.

In general, the paper emphasises the importance of ›local generalisations‹ and of seeing language as a ›system‹ of low-level regularities, not all-encompassing rules. While type frequency certainly seems important for productivity, inflectional morphology is a complex matter; productivity is also influenced by various factors of a more structural nature.

I. Introduction

Productivity is a notorious riddle in linguistics (see e.g. Bauer 2001) and the problem is perhaps particularly pressing in diachrony and psycholinguistics. In diachronic morphology, the concept of productivity is often invoked in discussions of analogical change in inflection in particular.

A widespread view is that type frequent inflection classes, i.e. inflection classes with many members, tend to gain new members, classes with few members tend to lose them. Sometimes, this view is described informally as ›the rich get richer‹, or ›the Matthew effect‹, in allusion to Matthew 25, 29 – »For to every one who has will more be given, and he will have abundance; but from him who has not, even what he has will be taken away«. In a textbook on analogy, Fertig ([2013], p. 116) says that »[f]or most [my emphasis; H. E.] analogical change, [...] the Matthew effect seems to matter much more than anything else«. This much is probably uncontroversial.

A stronger and much more controversial claim is that type frequency is the only factor deciding productivity, in other words that productivity is a direct consequence of type frequency, and that nothing else matters. This assumption is widespread (at least according to one of the referees for this journal); for example, it may seem to be the claim of Croft/Cruse ([2004], p. 296, 327). Coming from psycholinguistics rather than diachrony, Charles Yang (2016, 2018) is a recent, influential and unambiguous advocate of this view; he launches his »Tolerance Principle« (henceforth TP) as a formalisation of this strong and controversial claim (see Section II). Yang does not only assume that type frequency determines productivity; he also assumes that we can quantify exactly when a rule is productive or not, so that productivity (or lack thereof) can be predicted. The TP, to be detailed in section II.1 below, is an interesting and ambitious effort to say exactly how rich the rich have to be in order to get richer. Yang certainly deserves credit for taking the bull by the horns.

This paper is, however, a sceptic's response. At least from the point of view of diachronic morphology, the TP does not solve all our problems. The claim of Section II.2 is that two of Yang's most central examples, English verbs and German gender, are inadequately analysed. Section II.3 presents some conceptual concerns. The bulk of the paper, Section III, presents examples where the idea that productivity is determined by type frequency alone simply fails, so that in these examples, the TP is unhelpful. Admittedly, Yang (2016) does not restrict himself to inflectional morphology, but the bulk of his monograph is devoted to it, and we shall therefore only consider inflection here.¹ Section III.1 presents a Frisian and a North Germanic (Scandinavian) example of suppletion spreading, Section III.2 North Germanic examples of minority rules spreading. Both kinds of examples

¹ Also outside of inflection, one may doubt the idea that type frequency is all. For example, Goldberg ([2006], p. 99) argues for syntax that »learners do not generate new instances on the basis of type frequency alone«; she sees type frequency as one out of four factors that may be relevant for a pattern's productivity (p. 53). See also Gries (2018) and Wittenberg/Jackendoff (2018) for criticism of the TP that is not based on inflectional evidence.

are explicitly outlawed by Yang. Section III.3 presents, more briefly, examples in support from Latin, Romance and Slavic.

There is no denying that the Matthew effect usually is relevant for productivity, but other factors can come into play, so that the TP is too strong and makes incorrect predictions. From time to time ›marginal‹ patterns do spread; it does happen that the poor get rich. Relatively speaking, such cases may be rare (in inflection as in real life), but in absolute numbers, they are numerous enough to warrant our attention. They are also important enough. Even weak verbs in Germanic started out as an innovation (cf. e. g. Krahe/Meid 1969), and it is hard to believe that all of a sudden, weak inflection was the dominant pattern in the Germanic verbs. While type frequency certainly is important for productivity, it is not alone decisive always. We shall see that a number of supplementary factors may come into play. Thus, the point of this paper is not ›only‹ negative, to argue against an influential monograph; it is also to survey factors other than type frequency that are relevant for productivity (see Section IV.2).

The main claim of this paper, then, is that what begins as a lexically restricted pattern, a ›minor rule‹, can spread. Admittedly, this does not necessarily mean that the TP is entirely wrong; ›not always right‹ does not necessarily mean ›always wrong‹. However, I shall not address this larger issue.

II. A sketch of the TP

The TP is a somewhat intricate proposal, and not all aspects can be dealt with here. The TP is also primarily a psycholinguist's proposal, whereas this paper is primarily based on diachronic evidence. However, while Yang's starting point is psycholinguistics and language acquisition, in practice, he takes it for granted that diachronic evidence ultimately will give the same net result as evidence from acquisition (see e. g. II.2 below). This is certainly not unique to Yang; for example, coming from the opposite side, Wurzel (1984) focused on diachrony, and took it for granted that evidence from child language would converge. Yet if languages can be changed also by adult speakers, then there is no guarantee that what holds for child language, necessarily must hold for historical linguistics as well. We may hope that evidence from the two domains, diachrony and psycholinguistics, will converge, but we cannot be sure. Fertig ([2013], p. 24) states this view rather bluntly: »Children innovate in all kinds of ways in the early stages of acquisition. Much of this appears to be of little direct relevance to language change«.

If we are interested in language as a social entity, diachronic evidence is at least as relevant as evidence from child language, and if a number of changes

point in a certain direction, this fact may indicate some ›psychological plausibility‹ for a particular generalisation. Diachronic data can reasonably be expected to corroborate, and not to contradict, postulations about synchronic structure (cf. Maiden 2001, p. 45, see also Maiden 2018, p. 1, and Wurzel 2000). Diachronic data, this paper argues, indicate very clearly that ›minority patterns‹ can spread.

II.1 The program

At the outset of his monograph, Yang ([2016], p. 10–11) acknowledges that »a rule's quantitative coverage does not guarantee generalization« and that »an overwhelming statistical advantage does not necessarily translate into productivity«. Later in the book, however, this insight seems to be lost. At least, the claim is that type frequency (and not token frequency) »plays a decisive [my emphasis; H. E.] role in the acquisition of productivity« (p. 68). The claim is, furthermore, that »[m]inority productive rules are simply impossible« (p. 122). Yang is, in his own words, »tempted to take« the »strongest position one can take«, to »deny lexical analogy in the absence of productivity«, and the claim is correspondingly that »lexical/nonproductive patterns are almost [sic!] never overgeneralized« (p. 157). It is therefore consistent that Yang wants to »do away with notions such as ›semiproductivity‹ that one occasionally finds in the theoretical and psycholinguistic literature« (p. 39), and that he claims that »the distinction between productive and unproductive aspects of morphology appears completely [my emphasis; H. E.] categorical in child language« (p. 38).

The TP is a surprisingly exact formula (p. 9); productivity emerges directly as a mathematical function of type-frequency. I leave out the details here, but it is important to note that by the formula, there will be an exact tipping-point for productivity (so that, say, a pattern of six members will not be productive, while a pattern of eight can).

According to Yang ([2016], p. 9), children »consider a rule R in their language and evaluate its productivity according to the associated numerical value [...] the number of items to which the rule is applicable, and the number of items that defy the rule. The rule is accepted if »e« [the number of exceptions; H. E.] is sufficiently small; otherwise learners formulate a revised rule«. Thus, »the quantitative accumulation of exceptions can lead to the qualitative change in the productivity of rules«. This means that »the price of linguistic productivity arises from the quantitative considerations of rules and exceptions« and that »for a rule to be productive, the number of exceptions must fall below a critical threshold« (ibid.). (This threshold is defined by the mathematical formula.) »The unambiguous evidence for the acquisition of productivity is [...] overregularization« (p. 92).

However, while »the number of exceptions must be relatively low to guarantee the productivity of the rule«, Yang argues, interestingly, that »smaller« rules, those defined over relatively few items, can tolerate a relatively high number of exceptions« (p. 66).

An obviously controversial part of the argument is that children should maximise productivity, i. e., pursue rules that maximise productivity (p. 72); they should consider whether keeping a rule is the globally most efficient strategy.² This part of Yang's approach clearly conflicts with e. g. Joseph's (2011) view of generalisations, according to which speakers generalise ›locally‹; »speakers can focus on just a small amount of data [...] at any one time« (Joseph 2011, p. 415).

The present paper is not written from a generative point of view; Yang's book is. However, the ›global vs. local‹ disagreement is – at least these days – not necessarily a discussion between a generative and a non-generative view, since a generative scholar such as Westergaard (2009, 2014) is also skeptical towards the idea that speakers should opt for global efficiency. She argues in favour of ›micro-cues‹.³

II.2 Two examples that do not show what they were meant to show

Unfortunately, some of Yang's original examples in favour of the TP, viz. English verbs and German gender, do not stand up to closer scrutiny; this section will show why.

II.2.1 The never-ending story of English verbs

A classic in discussions of productivity and regularity is the case of English verbs. In discussing this issue, Yang ([2016], p. 28) enters the domain of diachrony, and in line with his overall argument, he claims that »overirregularizations« »turn out

² If the number of exceptions is »sufficiently large«, this may, according to Yang ([2016], p. 61) »slow down the overall time complexity to the point where resorting to full listing is more efficient; the Tolerance Principle provides a precise solution for what the tipping point may be«.

³ Conversely, also within Construction Grammar, the right level of generalisation has been debated; see e. g. Boas (2008) for some discussion. Also within current research on language acquisition, there has been considerable discussion on whether generalisations occur ›locally‹ or speakers store abstract patterns more ›globally‹ or both (see e. g. Ambridge 2020a and comments, including Ambridge's own partial retraction 2020b).

to be virtually absent«. Similarly, Yang ([2016], p. 34) talks of »the drastically [my emphasis; H. E.] different rates of overregularization and overirregularization«. This is debatable. A specialist on diachrony, Fertig ([2013], pp. 55f.), says that »[a]nalogical developments that can be described as irregularizations are considerably more common than textbook overviews generally indicate« and that »the discrepancy between change in the two directions is not always as large as it is claimed to be« (p. 80).

Yang ([2016], p. 158) argues that »the prima facie evidence against the productivity-based approach to change would be cases where an item has shifted from an undoubtedly productive rule to an undoubtedly nonproductive rule/pattern«. The terminology here is unfortunate, since a normal criterion for a pattern being productive is that it attracts new members, but the idea is clear enough.

A detailed examination of the English verbs is beyond this paper, but a couple of points are relevant. Firstly, also Yang acknowledges some productivity for the English strong inflection: »when *werede* became *wore*, the *bear–bore* pattern (strong class IV) must have been productive such that it could assimilate words that fit its structural description« (p. 157). This observation is surely correct, but the problem is that it is correct by definition. The productivity of strong class IV (*bear–bore*, *tear–tore*) does not explain the transition from *werede* to *wore*; the productivity of strong class IV, which is exactly what *werede* > *wore* bears witness to, needs explanation itself. Such an explanation may certainly involve the TP, but it is striking that Yang makes no attempt to show that it does. At least for the time being, the transition shows the productivity of a numerically minor pattern.

Outside of *wear*, Yang accepts that some examples of a change from weak to strong have been brought forward also in English. Yet he goes to great lengths to dismiss them, stating that the »most detailed study of this type is a quantitative analysis by Anderwald (2013), who examined the usage trajectories of [...] *dive-dove*, *plead-pled*, *drag-drug* and *sneak-snuck* in the Corpus of Historical American English«. What is not said here is that Anderwald ([2013], p. 148) also mentions »mildly irregular verbs like *burnt*, *learnt*, *spilt*, *spoilt*, etc.«. According to Anderwald, »there are some indications that [...] the irregular forms [in American English; H. E.] are a more recent development, or perhaps have been revived«. They are thus a potential problem for the TP, but Yang does not address them.

However, Anderwald's (2013) main focus is admittedly on *dive-dove*, *plead-pled*, *drag-drug*, *sneak-snuck* (and one verb that is irrelevant for present concerns). Yang ([2016], p. 158) argues that »three of the four counterexamples are only apparent. Two of these verbs (*dove* and *pled*) had variable strong and weak forms from the beginning of the corpus [i. e. around 1800]; the morphological changes in these cases have been a matter of frequency fluctuation«. It is far from clear how this argument can save the TP, since *dove* and *pled* remain innovations;

they must have originated at some point, and, again, an explanation of the productivity compatible with the TP remains to be seen. At least for the time being, such examples are an embarrassment for the idea that type frequency alone determines productivity.

Perhaps even worse, according to Anderwald ([2013], p. 155), the only source mentioned by Yang for this verb, the »new irregular form *dove* [in American English; H. E.] is formed in analogy with other lexemes rather than as a revival of a historical strong form«. Also for *pled*, Anderwald (p. 157) argues that the spreading is »very much a recent development«. As for *drug*, Yang is clearly right in calling its frequency »very low« compared to *dragged*, but that does not affect the correctness of Anderwald's observation that it »can be classified as a new irregular form«. ⁴

Finally, Yang's argument for dismissing *snuck* is that it conceivably *may* have become popular because of one particular author. Such a scenario seems somewhat unlikely (at least to me), and it should be contrasted with Anderwald's conclusion from her corpus study, which is that, for all the verbs, »we are most likely dealing with genuine changes from below« (p. 146).

In denying the productivity of *snuck* and the three others, Yang is denying the productivity of a central example of ›product-oriented‹ generalisations (cf. Bybee/Moder 1983).⁵ Simplified, a process-oriented generalisation has the format »in order to get to Z, take X and add Y«. A product-oriented generalisation, by contrast, has the format »Z will look like this«. To take an example, a process-oriented generalisation may be that »to get the English past tense, add -(e)d to the verb stem«, a product-oriented generalisation may be that »the English past tense typically ends in -(e)d«. There is an empirical difference; for example, only the product-oriented generalisation will cover both *killed* and *hid*.

Now, it is not surprising that Yang wants to dismiss *snuck*. In as much as the TP model appears to be all source-oriented, this is logical and consistent. However, there are many examples of such product-oriented generalisations; they cannot all plausibly be rejected (see III.1.2).

⁴ Also, if there are varieties of American English where *drug* is systematic, the problem for the TP remains; Yang does not tell us whether he has checked. It seems unlikely that the corpus should reflect mathematically the knowledge of each speaker.

⁵ For more discussion of product-oriented generalisations and the difference to process-oriented generalisations, see e.g. Bybee/Moder's (1983) classic paper or Köpcke/Wecker's (2017) recent defence of the need for both in German.

II.2.2 German gender and plural suffixes

Yang ([2016], p. 26) correctly points out that »The impoverished morphology of English inflections hardly seems appropriate to appreciate the complexity of languages«, and he does not base his theory only on English verb inflection. He takes on the challenges from German plurals; they are notoriously difficult to reduce to a binary regular/irregular distinction, since there are so many different plural suffixes, and it is not obvious that one of them is the only regular suffix. Yang accepts the challenge, claiming (cf. II.1) that »[m]inority productive rules are simply impossible under the Tolerance Principle« (p. 122). After a lengthy discussion, he concludes that the TP can be maintained, that »there is order in the apparent chaos of German plurals after all. But the horrors of the German language haven't exactly gone away: the child now just has to learn the marking of gender« (p. 136).

If this argument is accepted, the problem with the inflection classes no longer needs to bother adherents of the TP. However, there is a price for this solution. Unlike the linguist, children cannot specialise in either inflection class or gender; they must learn both. If Yang's account is on the right track, children cannot utilise inflection class information in order to learn gender. While this may be satisfactory from the perspective of maintaining the TP, it aggravates the problem of learning German gender (surely not an easy task anyway).

A priori, one possible cue for learning German gender might be the plurals, and three decades ago, Corbett ([1991], p. 49) suggested that it makes sense to take information about declensions as basic when learning gender. There are more inflection classes than genders in German, so it is descriptively easier to go from inflection class to gender than the other way around.

Furthermore, one may wonder why learners should stick to just one alternative, i. e., either only use genders to predict plurals or only use plurals to predict genders. Learners/speakers are generally opportunistic, clutching at every straw. (For psycholinguistic arguments, see Dąbrowska 2004; for diachronic arguments e. g. Maiden 2018, p. 317 et passim). So, why could learners not use a bit of both kinds of information? If a particular noun occurs very often in the plural, for example, it would make sense for learners of German to use their knowledge of the plural declension to infer the gender. In a psycholinguistic study of German gender, Müller ([2000], p. 352) says »the grammatical features gender and number are discovered simultaneously in language acquisition«.

Conceivably, the direction between learning plurals and learning gender is not unidirectional, so that sometimes, the plural can serve as a cue to gender, sometimes the other way around. There is diachronic evidence in favour of this position for German's relative Norwegian (Enger 2004). Thus, in late Old Norse

(Norwegian-Icelandic, after approx. 1200), the word for ›Easter‹, *páskar*, which occurred in the plural only, had its gender changed from feminine to masculine. This is presumably due to the ending *-ar*, most often associated with masculines. By contrast, the word for ›belt‹, *reim*, has had its ending changed in many dialects later on; the old suffix *-ar* (or a descendant) has been replaced by */er/*, the suffix typically found with feminines – presumably because *reim* already was a feminine. Norwegian is not isolated in this respect; Doleschal (2000) emphasises the possibility of bidirectional predictability between gender and inflection class for Russian. For the Low German dialects of East Friesland, Kürschner/Nübling ([2011], p. 378) suggest that »each gender corresponds with several declensions and gender can by no means be securely predicted from declension information«, so that it is difficult to conclude either way.

Admittedly, the East Friesland varieties are two-gender systems and thus different from Standard German, which is the variety Yang discusses. Yet also for Standard German, the prospects for the TP are bleak. For illustration, consider a very small snippet of the German noun inflection. Traditionally, inflection classes are posited on the basis of two principal parts, the genitive singular and the nominative plural (as noted by e. g. Blevins 2016, p. 187). Tab. 1, and the argument, is based mainly on Fedden/Corbett (2020), but see also Fedden [et al.] (2022).

Tab. 1: A very small piece of German inflection

	Fem.	Masc.	Masc.	Masc.
	›bee‹	›witness‹	›day‹	›knob‹
Nom. sg.	<i>Biene</i>	<i>Zeuge</i>	<i>Tag</i>	<i>Knauf</i>
Acc. sg.	<i>Biene</i>	<i>Zeugen</i>	<i>Tag</i>	<i>Knauf</i>
Gen. sg.	<i>Biene</i>	<i>Zeugen</i>	<i>Tages</i>	<i>Knaufes</i>
Dat. sg.	<i>Biene</i>	<i>Zeugen</i>	<i>Tag</i>	<i>Knauf</i>
Nom. pl.	<i>Bienen</i>	<i>Zeugen</i>	<i>Tage</i>	<i>Knäufe</i>
Acc. pl.	<i>Bienen</i>	<i>Zeugen</i>	<i>Tage</i>	<i>Knäufe</i>
Gen. pl.	<i>Bienen</i>	<i>Zeugen</i>	<i>Tage</i>	<i>Knäufe</i>
Dat. pl.	<i>Bienen</i>	<i>Zeugen</i>	<i>Tagen</i>	<i>Knäufes</i>

A noun such as *Knauf* ›knob‹, with *-es* in the genitive singular and umlaut in the nominative plural, cannot be a feminine; more than 90 % of such nouns are masculine. Of nouns inflecting like *Biene* ›bee‹, more than 90 % will be feminine, according to Fedden/Corbett. Do we really want to exclude the possibility that children can observe this? Fedden/Corbett also argue that, for the five largest

inflection classes in the German lexicon, the best prediction that can be made on the basis of gender accounts for less than 30 % of the data. The consequences for the TP should be clear.⁶

II.3 Two conceptual concerns

There is ample evidence for children being statistical learners, and (at least in my view) much of the task of learning a grammar has to do with probabilistic learning. However, the level of mathematical sophistication and exactness built into the TP is somewhat surprising, even for an old adherent of probabilistic learning. Yang ([2016], p. 64) makes a similar point very directly: »At the present time, it is not clear how the Tolerance Principle is executed as a cognitive mechanism of learning; surely children doesn't [sic!] use calculators.« This cavalier attitude towards the cognitive side seems somewhat worrying, especially when it comes from a psycholinguist. Apart from that, it is hard to believe that language acquisition should be quite so mathematical (cf. also Wittenberg/Jackendoff 2018, p. 790). Since Yang clearly sees his TP approach as being within the Chomskyan mould, one wonders what happened to the old generative maxim that grammar does not ›count‹ in the arithmetical sense.⁷

⁶ Recently, Köpcke/Wecker (2017) have presented arguments why product-oriented and source-oriented generalisations are both needed for German plurals; since Yang's model excludes product-oriented generalisations (cf. section III.1.2), this is also a problem.

⁷ On a historiographic note, it is interesting that in the 1990s, staunch defenders of Universal Grammar would argue that type-frequency did not alone determine productivity and default status (e. g. Marcus [et al.] 1995). Twenty years later, Yang argues – while paying homage to the same research program – that type-frequency determines productivity completely. UG remains as a basic premise, while at the same time largely undefined: »These general constraints on language, whatever final form they turn out to take, appear universal and inviolable« (Yang 2016, p. 218, my emphasis; H. E.). Several scholars have asked exactly what the empirical content of the UG program is (e. g. Tomasello 2004, Dąbrowska 2015). Yang's carefree attitude here can hardly calm them. Yang is happy to equate UG/LAD with general learning mechanisms: »Conceivably, the so-called Language Acquisition Device [...] will no longer be necessary as a specialized, and independently evolved, module of the mind/brain« (p. 215). One wonders what is left of UG, in the earlier sense of cognitive principles that are language-specific. No functionalist (or even behaviourist for that matter) has ever denied that humans are born with a general capacity for learning (cf. Dąbrowska 2015). My point is not that there is anything wrong in being generative or an adherent of the UG hypothesis, only that the empirical content of the hypothesis has been very flexible. Also adherents of the UG program, such as Lasnik/Lohndal ([2010], p. 46–7) note that the research goal now is to »reduce the principles of UG to their ›barest‹ essentials and seek principles that are more general in nature, e. g., as part of our general cognition or even of biological systems more generally«. Golumbia ([2010], p. 29) says ›Minimalism

The TP is based on the Elsewhere Condition. According to Yang (p. 50), the key claim of the Elsewhere Condition is that the computation of rules and exceptions is serial. (Before uttering an inflected form of a certain word, then, speakers presumably first have to check that this word is not irregular.) To assume as a basic premise that language computation is exclusively serial is probably to build a theoretical house on psycholinguistic sand. Kapatsinski ([2018], p. 738) notes that the assumption of serial search has never been accepted in the field of psycholinguistics; he argues that the assumption is incompatible with the parallel nature of the brain, and that »[w]ithout serial search, there is no motivation for the TP«.

III. Counter-examples

The objections raised in Section II.3 attest to a theoretical starting-point that differs from Yang's, so they are unlikely to convince anybody that does not already agree with my starting-point. In other words, if the proof of the proverbial pudding really is in the eating, as emphasised also by Yang (p. 77), these objections might seem less important. If the empirical predictions of the TP really work, type frequency alone determines productivity. Unfortunately, the predictions do not seem correct; at least, they did not seem entirely right for two of Yang's own examples (II.2), and the claim of this section is that they do not hold for a series of other Germanic examples, either.

Recall that »[m]inority productive rules are simply impossible under the Tolerance Principle« (Yang 2016, p. 122). This claim is too strong. »Lexical rules« or »minority rules« can be the basis for analogies; they can become productive. Yang (p. 92) emphasises that the »unambiguous evidence for the acquisition of productivity is [...] overregularization«. Therefore, we shall look at some cases where what starts as a very restricted pattern spreads. Section III.1 presents two Germanic cases of suppletion spreading, III.2 four North Germanic cases of minority rules being productive. III.3 presents some parallels from outside of Germanic.

is Functionalism«, and that the Minimalist Program in some ways is »remarkably close to the perspectives offered by Chomsky's most prominent [...] opponents«.

III.1 Suppletion spreading

Productivity on the basis of one example is certainly not expected under any theory, but under the TP, it is explicitly outlawed – as it must be, if type frequency is supposed to be the sole determinant of productivity. Commenting on such English verbs as *say – said*, *come – came*, Yang ([2016], p. 82, footnote 4) says the rules instantiated »cover only one verb [...] are in effect suppletive [...] and have no potential for generalization at all«. Similarly, minority rules should be impossible. This is truly logical and consistent. If only one verb adheres to a pattern, this pattern has to be lexically listed, and like many other theorists, Yang tries to keep lexicon and grammar apart. Recall now that »[t]he unambiguous evidence for the acquisition of productivity is [...] overregularization« (p. 92). If suppletion can spread, that is at least problematic for this view. Below, I present two cases, from Frisian and Scandinavian. Both indicate that suppletion actually can spread, even if the former is more clear-cut.

III.1.1 The clear-cut case: Frisian

Nübling ([2000], p. 205, 228) presents an interesting case of suppletion spreading. The Frisian verb *jaan* »give« has had its past tense changed from *jef* to *joech*. The change is, according to Nübling, due to analogy with *sloech*, the past tense of *slaan* »hit«. What makes the case of *jef* > *joech* so interesting is that prior to the change, *slaan-sloech* was an isolated anomaly, a suppletive pattern. It is admittedly not often that suppletion spreads, but by the TP, it simply cannot. This strong claim is clearly not supported by the Frisian facts.⁸

III.1.2 The less clear case: Scandinavian

The past tense of *stå* »stand«, in some few Norwegian dialects, is /stu:g/; the final /g/ is historically unexpected. Venås ([1967], p. 289) explains the innovation as analogy due to /dru:g/ and /gnu:g/, the past tense of *dra/draga* »pull« and

⁸ An anonymous reviewer asks whether it is really justified to talk of productivity if an inflectional pattern is transferred to one single verb. Following Wurzel (1984) and Yang, I take a pattern to display productivity if it acquires new members. The reviewer goes on to suggest that this change could rather be called a local analogy. I beg to differ. The widespread distinction between analogies and rules is certainly not theory-independent, and given that Yang argues that suppletion cannot be generalised at all, this is simply a counter-example.

gnaga ›chew‹ respectively. That is entirely plausible, but, importantly, these verbs differ significantly from *stå* in their infinitive and present tense forms. Compare the infinitive /sto:/ ›stand‹ vs. /dra:(ga)/ ›pull‹, /gna:ga/ ›chew‹.

The innovation /stu:g/ is not unique. We find /lu:g/ ›laughed‹ in a larger area than /stu:g/; the analogical mechanism is presumably the same, but the infinitive and present tense stem differs yet again, cf. /le:/, *le*. In some areas, we also find /du:g/ as the past tense of *dø* /dø:/ ›die‹. In the three verbs ›stand‹, ›laugh‹, ›die‹, the /g/ cannot be due to sound change, and it is traditionally accounted for morphologically (Venås 1967).

It may – fairly – be objected that by the strictest definitions, this is not suppletion spreading, since there were two verbs (*draga*, *gnaga*) that ended in /u:g/ in the past tense at the outset, and not only one. However, it is not the pattern (the rule) that is spreading. In order to get from /dra:/ to /dru:g/ and from /gna:ga/ to /gnu:g/, we replace /a:(ga)/ by /u:g/. Yet in order to get from /sto:/ to /stu:g/, we replace /o:/ by /u:g/; in order to get from /le:/ to /lu:g/, we replace /e:/ by /u:g/ and in order to get from /dø:/ to /du:g/, we replace /ø:/ by /u:g/. To account for these five verbs, then, four different rules are necessary, even if the past tense forms are so similar they rhyme. In his study of Swedish morphology, Kiefer ([1970], p. 124) explicitly labels the cognates *log*, *dog* suppletive. This is understandable, given that the ›base forms‹ and the processes differ. Yet it is unsatisfactory, because the similarity between the past tense forms is not captured. Since speakers by analogy have made the verbs more similar than they were, linguists should be able to capture this. And the similarity is easily described by means of a product-oriented generalisation (cf. II.2.1). Informally, there are some verbs that in the past tense end in /u:g/.

However, such a move would violate the spirit behind the TP, which is a strictly source/process-oriented approach. Yang's model does not open for product-oriented generalisations. Yet if one insists on excluding product-oriented generalisations, Norwegian /stu:g/ etc. are suppletive, and thus a problem for the TP.⁹

III.1.3 Why suppletion can spread

A priori, it may seem strange to talk about ›generalisations on the basis of one exemplar‹, and so many linguists (not only Yang!) would hesitate to accept cases

⁹ If product-oriented generalisations are accepted in the TP framework, this counter-example evaporates. In that case, however, one will have to accept semi-productivity, a notion that Yang ([2016], p. 39) is eager to dispose of. (The past tense of *be* ›beg‹ /be:/, /ba:/, has not yet shown any sign of being replaced by /bu:g/, for example.)

like the Frisian one in III.1.1. Yet humans often reason along other lines than those prescribed by the laws of logic. For example, they can reason abductively (cf. Janda 1996, Andersen 1973) – and they quite often do so.

Let us suppose that we live in Northern Europe and, like a number of other inhabitants (unfortunately!), harbour unreasonable prejudices against, say, refugees. We may know a number of perfectly nice and respectable refugees, but we still cling to our stereotypes on the basis of the one less respectable refugee that we have heard of; we insist that, at heart, the nice and respectable ones are all ›exactly like the bad guy‹.

This is a ›generalisation based on one item‹. People can make generalisations based on one example even in the face of overwhelming evidence to the contrary. An obvious question is under what circumstances they do so. We need not pursue the analogy with prejudices in Europe any longer, but we shall return to the morphological version of this question (when do small patterns become larger?) in Section IV.2 below. On a more linguistic note, the idea that a single word in some respects is also a rule is found in frameworks as different as Lexical Phonology (e.g. Kenstowicz 1994, ch. 5) and Construction Morphology (Booij 2010). For present purposes, the relevant point is made by Maiden ([2018], p. 302): »learning distributional patterns and abstracting from them are in principle different operations. An idiosyncratic pattern may be learned for a single lexeme [...] without speakers necessarily [emphasis original; H. E.] transcending the individual cases and making abstract generalizations« – by implication, they also can, as Maiden shows.

The spreading of suppletion is clearly not what anybody would predict. Yet the problem becomes less pressing if we do not insist on a strict dissociation between lexicon and grammar – and we probably should not insist on that, anyway, cf. e.g. Jackendoff/Audring (2018, p. 395, 398 and references). Giving up the strict split is incompatible with the spirit of the TP, though.

III.2 Minority rules can spread I: North Germanic evidence

III.2.1 North Gudbrandsdalen (Norwegian) /de/

In the dialect of North Gudbrandsdalen, Norway, /de/ is the regular definite suffix in the singular of neuters that end in /e/ in the indefinite. Thus, /ju:ɾe/ ›field, acre‹ has as its definite singular /ju:ɾede/, /æpɾe/ ›apple‹ has as its indefinite singular /æpɾede/. This feature of the dialect will often catch the ear of other speakers of Norwegian, who find this suffix highly unusual. Compare e.g. Oslo, where the corresponding definite singulars would be /ju:ɾe/, /epɾe/ respectively

(with no suffix/a zero-operation; i. e., for this class of neuters, the definite singular is syncretic with the indefinite, unlike other neuters). However, the suffix /de/ has also intrigued linguists, since it does not follow from any ›sound law‹. Its origin is not entirely certain, but the best suggestion around is that it originates in a very few neuters that ended in an Old Norse *ð*, such as *høfuð* ›head‹. The story involves re-segmentation (or ›secretion‹, in Haspelmath's 1995 terminology). Old Norse *ð* becomes *d* in this dialect (unlike Oslo and most other dialects, where it is simply lost), and the consonant is later reanalysed as part of the suffix rather than of the stem, perhaps in part because of onset maximisation. In other words, *høfuð-it* is reanalysed as *høfu-dit*. A number of phonological changes that need not concern us now also take place, and today, the definite singular is /hugude/, and the indefinite is /hugū/.

Apart from *høfuð* and *herað* ›district‹, I cannot think of any bisyllabic neuter nouns in Old Norse ending in *ð*. By contrast, there are many neuters ending in an unstressed vowel, such as Old Norse *epli* ›apple‹; Conzett's ([2007], p. 33) estimate is more than 400. Yet the suffix /de/ has spread to bisyllabic neuters that end in an unstressed vowel, such as *epli*. Thus, the case indicates that analogy can work from a minority of no more than a handful, influencing a majority of 400. This is completely incompatible with the TP. The case indicates very clearly that type frequency cannot be a strict prerequisite for the generalisation of morphological patterns (Enger 2007a, p. 57).

III.2.2 Stavanger /o/

In Old Norse, there are two main classes of feminines, traditionally referred to as ›strong‹ and ›weak‹. The former end in a consonant in the indefinite singular, the latter end in an *-a*. Examples of the former include *sól* ›sun‹, of the latter *vísa* ›song‹. In the definite singular, the former end in *-in*, the latter in *-an*, cf. *sólin* ›the sun‹ vs. *vísan* ›the song‹. The strong feminines are numerically predominant in Old Norse. Beito ([1954], p. 15) claims there to have been 60 % strong feminines, 40 % weak, while Conzett ([2007], p. 33), who presents a more fine-grained classification and somewhat different figures, assumes 45 % weak feminines, 55 % strong. On either count, the difference is not overwhelming, but it is there.

However, there are dialects of Norwegian today where it is quite clear that the old definite singular suffix of the weak feminines (the minority) has been extended to the strong feminines (the majority). This holds, for example, for the dialect of Stavanger, according to traditional analyses (e. g. Berntsen/Larsen 1924/1978, p. 10). In this dialect, all feminines end in /o/ in the definite singular, so that /su:lɔ/ ›the sun‹ (›strong‹) has the same suffix as /vi:so/ ›the song‹ (›weak‹), and

the /o/ in the strong noun cannot plausibly come from Old Norse *-in* by ›sound law‹; it probably has developed from *-an*. The minority suffix has ousted the majority suffix.

III.2.3 A smaller weak class in Norwegian and Swedish verbs

In Old Norse, there are strong and weak verbs. The weak ones are usually subdivided into three classes, based on affixes. Tab. 2 shows the affixal conjugation of Old Norse, based on Noreen 1923, Haugen 1995:

Tab. 2: Affixal inflection classes for Old Norse verbs, 3. sg.

	strong	<i>telja</i> ›count‹	<i>kasta</i> ›throw‹	<i>duga</i> ›suffice‹
Present	-r	-r	-ar	-ir
Past	-∅	-di	-aði	-ði
Supine (Past participle)	-it	-t	-at	-at

It is uncontroversial that the *kasta* class is the largest of the weak classes; it has more members than the others. However, in many dialects in Trøndelag, Norway, the suffix from the *duga* pattern of Old Norse has gained ground at the expense of the suffix from the *kasta* pattern to the extent of replacing it entirely (cf. Dalen 1990, p. 134; Enger 2007b, pp. 295 f.). Thus, today, we can find forms such as /kasti/ in the present tense (the r-loss need not concern us here); one may say there are only two weak inflectional classes in these varieties. This is in all likelihood a morphological development (and not due to ›sound law‹).

Also in much West Norwegian, e. g. Stavanger, the present tense suffix originally associated with *duga* has been unexpectedly productive. As in Trøndelag, the development is not due to regular phonological change, but to morphological analogy. What makes Stavanger different from at least much of Trøndelag is that the change does not affect all members of the *kaste* conjugation. In other words, the change does not take place on the class level, rather, on the lexeme. In a TP perspective, a minority suffix spreading is clearly problematic, still, in both cases, as noted also by Lykke (2020).

The facts of unexpected productivity in Western Norwegian verb inflection are well known and have received much attention (e. g. Venås 1974, Wetås 2012). Fifteen years ago, Enger (2007b) suggested relating the spreading of *duga*

affixes to the No Blur Principle, NBP (see Carstairs-McCarthy 1994, 2010, Cameron-Faulkner/Carstairs-McCarthy 2000). The idea was that even if the *kasta* class was the largest in terms of type frequency, the spread of affixes from this class would have meant a violation of the No Blur Principle. Thus, there were structural restrictions on productivity that did not have to do with type frequency.¹⁰ At least, so the argument went.

There is much more to be said about the NBP, its different versions and many good sides (cf. especially Carstairs-McCarthy 2010), but what is relevant right now, is that recently it has been argued that the NBP is somewhat too strong. Certainly, far from all putative counter-examples stand up to scrutiny (cf. Carstairs-McCarthy's 2010 very persuasive arguments against a number of them), but some few do. Also, recent contributions suggest that the NBP may not be an independent principle, but rather a by-product of principles relating to inflectional entropy, see Ackerman/Malouf (2013, 2014), and Blevins (2015, 2016). Yet, even according to these critics, the general idea is probably on the right track, and this more general idea is not compatible with the TP or with the idea that type frequency determines productivity; there is more to productivity than mere type-frequency.

According to Venås ([1974], pp. 21 ff., 97), the pattern associated with *duga* has been more productive for verbs with a higher token frequency. If so, that is an additional problem for the TP, since token frequency is not given much room in that approach. Yang emphasises that it is type and not token frequency that decides productivity (cf. II.1 above); if Venås is right, however, token frequency can also play a role.

Swedish certainly does not develop out of Old Norse. Yet the historical details need not concern us here; the verb inflection in Old Swedish is similar enough to that of Tab. 2 that we can make two additional observations. In current Swedish, the *kasta*-class is the only fully productive conjugation, and it is numerically dominant. Yet in the dialect of Eskilstuna, the strong supine (roughly, past participle) suffix /i/ has spread to the *kasta*-class. This indicates that productivity does not follow automatically from type frequency.

In Standard Swedish, furthermore, the word tone associated with strong verbs has spread to the *duga* class, even if there were more *duga* verbs than strong ones. There are probably structural reasons for this development (cf. Enger 2014 for some suggestions), but the development is hard to reconcile with the TP, or with the idea that type frequency alone determines productivity.

¹⁰ The details are fairly technical and not essential here (but see Enger 2007b).

III.2.4 On what words is a productive rule productive?

The TP is an attempt at making productivity a precise notion so that we can quantify first and then predict productivity, but the term ›productive‹ makes sense only if it operates on a certain class or set. An inflectional class X can only be productive for words (or morphemes, or stems, depending on framework) that have a certain characteristic Y, whether Y is ›being a verb‹, ›being a noun ending in an unstressed vowel‹ or whatever else the reader might think of. However, this raises a classic question, namely how the native speaker knows what the relevant class is (Bauer 1997, p. 554). This question is left open by Yang (2016), and so it is open to debate whether the predictions of the TP really are so precise.

An illustration of the problem comes from Wurzel ([1984], pp. 149 f.; 2000, pp. 351 f.). In Old Norse (aka Old Icelandic), there are (at least) three inflection classes for feminines, traditionally referred to as i-stems, \bar{o} -stems and consonantal stems. There were many feminine i-stems and \bar{o} -stems, very few consonantal stems. Yet, Wurzel ([2000], pp. 351 f.) notes that

»certain words from the i-declension like *ǫnd* ›duck‹ and *ǫlpt* ›swan‹ and certain words from the \bar{o} -declension like *hind* ›hind (of venison)‹ transferred to the consonantal declension more or less consistently already in Old Icelandic [...] This seems surprising at first glance. When one examines the facts more closely, however, it becomes clear that the class of consonantal feminines contains, among others, a large collection of animal names [...] while each of the other two classes contains only a few animal names. The interpretation of these class transfers has to do with speakers' taking the semantic property ›animal‹ as a criterion for the class membership of lexemes [...] the consonantal declension became the preferred class for feminine animals [sic!] ending in a consonant, and appropriate words of both competing classes joined this one. Thus speakers exploit not only the common, fundamental syntactic property of gender and the equally common, fundamental phonological property of word ending, but also the quite specific semantic property ›animal‹ for class specification, a property which (as far as anyone knows) had never played a role in Germanic and Nordic grammar before.«

Thus, Wurzel acknowledges that sometimes a minority rule can become productive, but, he argues, this is due to speakers having re-defined the classes, so that what looked like a minority – the consonantal inflection – is now a majority within a redefined set (feminine names of animals).¹¹

Obviously, this example illustrates a diachronic change; a semantic property ›name of animal‹ has become relevant in inflection. Equally obviously, while dia-

¹¹ Wurzel's starting-point was not terribly different from Yang's, in that Wurzel also believed that type-frequency would determine productivity, to an extent that made Maiden (1996) characterise his belief as ›sanguine‹. Unlike Yang, however, Wurzel explicitly opened for other factors possibly intervening. See IV.2 below.

chronic change is the bread and butter of the diachronic linguist, it is not so for the psycholinguist, and Yang is certainly excused for not having the possibility of diachronic change at the top of his priorities. Yet unless we have some way of knowing how speakers delimit their classes, progress is unlikely to be made here. The problem needs to be addressed by adherents of the TP.

III.3 Minority rules can spread II: Evidence from Romance and Slavic

In sections III.1 and III.2, we have looked at six examples of minority rules spreading, and they are certainly not unique. My main point is not revolutionary, either. For example, Joseph ([2011], p. 408) argues that »grammatical material with a general application can often be shown to arise from very particular combinations«. The Latin imperfect *-bā-*, »a formation found for all the verbs in all of the regular conjugation classes of Latin« started, according to Joseph, in a restricted part of the verbal system. Joseph ([2011], p. 409) goes on to say that »[e]xamples of this sort can be multiplied easily«, and indeed they can.

For Germanic, Dammel (2011) presents examples other than the ones here, showing that inflection classes are not only subject to *Abbau*, reduction, but also to *Aufbau*, introduction and spread, and such classes have to begin somewhere. None of Dammel's examples shows an innovation becoming a majority in one giant leap.

Outside of Germanic, we may look at the origins of the so-called L-morpheme in Romance. Compare Tab. 4:

Tab. 4: A Portuguese example of the Romance L-pattern, verb *ter* ›have‹

1. sg.	2. sg.	3. sg.	1. pl.	2. pl.	3. pl.	
<i>tenho</i>	<i>tens</i>	<i>tem</i>	<i>temos</i>	<i>tendes</i>	<i>têm</i>	Present indicative
<i>tenha</i>	<i>tenhas</i>	<i>tenha</i>	<i>tenhamos</i>	<i>tenhais</i>	<i>tenham</i>	Present subjunctive

(The shading is meant to show the similarity with an ›L‹, the sequence *nh* is pronounced /ɲ/.)

It is hard to see any good reason for this formal identity – or any natural class at work. Yet the identity in form has probably been noticed by speakers, since abundant diachronic evidence indicates productivity for the L-pattern. Independently of the particular material that may ›fill‹ the particular cells, there is a pattern of identity. So how did the pattern begin?

This is discussed at length by Maiden ([2018], ch. 5), but in what follows, I rely extensively on an earlier discussion (Maiden 1996), which is sufficient for present purposes. Maiden ([1996], p. 167) summarises the case as one in which »a small handful of Romance verbs acquired doubly disparate patterns of allomorphy in their root: the phonological details of the resulting alternants were often quite different from verb to verb, and characterized a disjunct and highly ›unnatural‹ set of morphosyntactic properties«. Maiden (pp. 167 f.) emphasises that »the relevant allomorphy originates in just a handful of verbs [...] yet it is diachronically productive«, and that »there can be no question of quantitative pressure« giving rise to the pattern »because the apparent historical basis for the analogy is restricted to a handful of verbs« (p. 193).¹² According to Maiden (p. 197), »the evidence of the Romance gerund suggests [...] that [...] local [...] and very rare, paradigmatic patterns have the potential to increase their domain«. This observation is in perfect agreement with Joseph's ([2011], p. 409) claim that »generally applicable material [...] starts out as highly particularized as to its originating context«, and with the examples in III.1 and III.2.1.

Also according to Janda ([1996], p. 1), whose focus is on Slavic, there are »cases where the progress of analogical change seems to buck the current, carrying an ›irregular‹ morpheme upstream to spawn«. She goes on to argue that

»The unexpected productivity of marginalized morphological forms is an under-studied phenomenon [...] On occasion a defunct morpheme *does* [emphasis original; H. E.] bounce back from the brink of extinction in a new productive role [...] In fact, this phenomenon is not all that rare; as we will see, it has occurred repeatedly in the histories of all the Slavic languages, which probably means that it happens in the histories of most other languages as well.«

One of her examples, simplified here, is that a general first person singular marker in the present tense, in current Macedonian, Slovene and Slovak, started out as being restricted to only five verbs in Old Church Slavonic.¹³

12 Maiden ([1996], p. 192) also refers to examples of metathesis in Tunis Arabic, as discussed by Kilani-Schoch, who »develops an example of a highly productive rule of deverbal derivation [...] consisting purely of metathesis which is unnatural in universal terms and does not correspond to any numerically pre-dominant pattern in the dialect, but none the less fits with general typological characteristics of Arabic«.

13 If additional examples are needed, here are two more from Romance. The very productive inflectional plural marker *-uri* in modern Romanian arises from the resegmentation of a handful (a dozen?) of Latin neuter plural nouns such as *tempo-a* ›times‹, reanalysed as *temp-ora* > *timp-uri*; today, we even find in Romanian *harddisk-uri*. In Romansh first person singular present tense forms, the ending *-el*, originally phonologically explicable in just a handful of verbs, has taken over the entire verb system.

III.4 Summing up

Minority rules and even suppletion can be productive. Evidence from North Germanic, West Germanic, Romance and Slavic support this claim.

IV. Conclusions

IV.1 A possible objection

The TP yields clear-cut and strong empirical predictions, and from the point of view of at least one influential philosopher of science, Karl Popper (1972), that is a virtue and a hallmark of true science. Weaker versions of the Matthew effect do not necessarily yield equally strong predictions. Thus, an objection at this stage may be: »What is the alternative hypothesis, exactly?«

To my mind, such an objection seems misguided. The reason is that, according to the TP, minority rules in inflection should never become productive, at least not by a morphological path. This claim is not only strong, it is also flatly wrong.¹⁴ Thus, clinging onto the claim does not seem reasonable, irrespective of philosophical ideals.

While the TP may be controversial, the Matthew effect is not, and it takes care of most examples (cf. section I), although not the ones presented in this paper. For them, we have to accept that a number of additional factors are also at play; some have already been mentioned, and I summarise them in section IV.2.

IV.2 Returning to the examples: other factors than token frequency are relevant

While this paper cannot present a full-fledged and detailed alternative to the TP, the general conclusion seems clear: Type-frequency is relevant for productivity in inflection, but it is not the only factor. A number of other factors, of a structural

¹⁴ The observation is far from new. In a textbook, Fertig ([2013], p. 113) points out that »there are a number of cases where a pattern (eventually) spreads significantly in spite of having once been restricted to a very small number of items [...] Especially interesting are cases where a pattern that originally occurred in just a single item comes to be highly productive.« Fertig mentions the suffix *-ess* as a case in point and refers Brugmann (1885) (!) for further examples.

kind, must supplement the Matthew effect, and indeed, such factors have been suggested in the literature already.

Within Natural Morphology as developed in the 1980s and 1990s, at least two such factors were accepted, viz. universal, semiotic naturalness and language-specific ›normalcy‹ (see e. g. Wurzel 1984, Gaeta 2018). While Natural Morphologists such as Wurzel (1984) would assume that one of these factors should consistently outrank the other, the relation is probably one of vacillation (Maiden 1996, p. 192). Sometimes, one wins, sometimes the other – and that alone may be sufficient reason to abandon the aim of precise prediction.

Furthermore, there are morphomic patterns that may over-rule type frequency, as Maiden (2018) has made abundantly clear. In his discussion of the Romance gerund, Maiden ([1996], p. 169) argued that »what is important in the propagation of phenomena which are ›unnatural‹ in system-independent terms is not quantitative predominance, but simply stability in the form – meaning relationship«. In later publications, Maiden (2013, 2018) has suggested that such stability and predictability is ultimately an important factor behind all matters morphomic.

A related strand of research has argued that sometimes, the influence of token frequency on productivity may be over-run by the No Blur Principle (for which see Carstairs-McCarthy 1994, 2010 and Cameron-Faulkner/Carstairs-McCarthy 2000). While this principle may have been somewhat too strong, the point of interest is that it worked so well (as Blevins 2016 points out).

A different tradition has emphasised that token-frequency has a role to play in morphological theory. It has been argued that what is the rule for low-frequency items is not necessarily so for high-frequency items; see Nübling (2000) and Dammel ([2011], pp. 277–280) for an extensive defence of this view. From a slightly different perspective, also Blevins ([2016], pp. 199–204) argues that irregularity is useful.

With these observations in mind, let us return to the examples above. The Frisian case of suppletion spreading (III.1.1) shows the role of token-frequency, that what is the rule for low-frequency items is not necessarily so for high-frequency items.

As for the unexpected productivity of *duga* affixes (III.2.3), if Venås's (1967) analysis is right, token frequency is relevant here too. However, if the account advanced by Enger (2007b) is on the right track, the explanation may involve the No Blur Principle of Carstairs-McCarthy (1994, 2010) or a later equivalent in terms of inflectional entropy. Neither of these accounts rely on type frequency.

The East Norwegian case of /stu:g/ etc. (III.1.2) indicates the need for product-oriented generalisations; if they are accepted, the example is unproblematic for morphological theory in general, but the example remains an embarrassment for the idea that type frequency is everything.

The North Gudbrandsdalen case of /de/ (III.2.1) spreading may, in Natural Morphology terms, be due to either semiotic naturalness (it may be better to have a clear-cut marker of definiteness than to have zero marking), ›normalcy‹ (the norm in North Gudbrandsdalen is that definiteness is expressed, morphologically), or both. The spreading of the weak feminine suffix in Stavanger (III.2.4) indicates that, when two groups are both sufficiently large and the difference between them small, it is not terribly surprising that the ›minority rule‹ should win. Why would speakers compute the difference in *e x a c t* numerical detail? Yang dismisses the question (II.3), and that is not good enough.

The reinterpretation of the consonantal class in Old Norse (III.2.4) illustrates the pertinence of Bauer's (1997) question: »how does the native speaker [...] know what the relevant [...] class [...] is?«. There is probably no final answer to that question. That is of course unsatisfactory, but at least, there is a good reason why: the notion of ›inflectional class‹ rests on the notion of ›inflectional similarity‹, which probably cannot be sharply delineated for all purposes.

As for the Romance and Slavic cases briefly mentioned in III.3, I have nothing original to add to the sources I rely on. Maiden (2018) now sees the Romance gerund in ›morphomic‹ terms, so even if it does not fit the TP perspective, or the idea that type frequency is everything, it fits a different one (also known as ›morphology by itself‹). Janda (1996) has shown how her Slavic examples fit into a perspective of humans as ›abductive‹ learners, in the sense of Andersen (1973), compare also the ›refugee example‹ in III.2.3.

IV.3 Envoi

The TP is an ambitious attempt at accounting for productivity, in a psycholinguistic perspective. It is also a very strong version of a widespread intuition that type frequency is relevant for productivity; by the TP, type frequency alone is all that matters for productivity. Thus, productivity emerges as a direct mathematical function of type frequency.

This paper has had a somewhat negative focus, in that I have brought forward a number of diachronic examples showing a simple point: minority patterns can become majority patterns,¹⁵ thus indicating that the TP is insufficient; type frequency alone cannot be the only factor triggering productivity in all the exam-

¹⁵ Another relevant observation is that major patterns can lose productivity. This observation has not been in focus here, but Kapatsinski ([2018], p. 740) rightly observes that it also is a problem for the TP, and that it is not unheard of in diachrony. It is of course also a problem for the idea that type frequency determines productivity.

ples. These examples may seem marginal. However, labelling a set of examples marginal will not make them go away; in the words of Janda ([1996], p. 6), »it would be dishonest to ignore data just because they are a bit unusual and appear theoretically unattractive«.

It is not really terribly surprising that inflectional morphology should not be all about frequency (be it type or token). A number of theorists have argued that languages are better seen as ›systems‹ of partly competing low-level regularities than as systems of all-encompassing, ›global‹ rules, and that we should focus on ›local generalisations‹ (Joseph 2011), or, in a different terminology, ›micro-cues‹ (Westergaard 2014). This means that inflection is subject to complex interaction between many different factors (see also e.g. Wurzel 1984; Carstairs-McCarthy 2008, 2010; Maiden 2018 for similar conceptions of morphology).

My take-away message is that type-frequency cannot be the only factor causing productivity. The reason is, in Joseph's ([2011], p. 415) words, that »speakers favour local solutions (small-scale generalizations)« and that »[e]ven full-scale generalizations start small«. If we accept the view that generalisations can be local in nature, it is not quite so strange that rather isolated minority patterns can spread. Inflectional morphology is subject to a number of competing factors, some of which are surveyed in IV.2, and it has a degree of autonomy of its own, so it would be surprising if type-frequency alone could decide productivity. Analogy can work on the basis of small groups, and a variety of structural and language-specific internal factors can intervene. Therefore, productivity need not be a direct consequence of relative type-frequency, even if it often is.

Article note: This article was written as part of the international research project ›MultiGender‹ at the Centre for Advanced Study at the Norwegian Academy of Science and Letters in Oslo during the academic year 2019–2020. I am grateful to the project and the centre for providing ample time and a supportive environment, and to Terje Lohndal, Martin Maiden and two anonymous reviewers for constructive and helpful comments.

Bibliography

- Ackerman, Farrell and Robert Malouf 2013: Morphological organization: the low conditional entropy conjecture, in: *Language* 89, p. 429–465.
- 2014: The No Blur Principle Effects as an Emergent Property of Language Systems, in: Hanna Jurgensen [et al.] (eds.): *BLS* 41, p. 1–14.
- Ambridge, Ben 2020a: Against stored abstractions. A radical exemplar model of language acquisition, in: *First Language Acquisition* 40, p. 509–559.

- 2020b: Abstractions made of exemplars, or ›You're all right, and I've changed my mind: Response to commentators, in: *First Language Acquisition* 40, p. 640–659.
- Andersen, Henning 1973: Abductive and deductive change, in: *Language* 49, p. 765–793.
- Anderwald, Lieselotte 2011: Norm vs variation in British English irregular verbs: the case of past tense *sang* vs. *sung*, in: *English Language and Linguistics* 15, p. 85–112.
- 2013: Natural language or prescriptive influence? Throve, dove, pled, drug and snuck in 19th-century American English, in: *English World-Wide* 34, p. 146–176.
- Bauer, Laurie 1997: A class of English irregular verbs, in: *English Studies* 78, p. 545–555.
- 2001: *Morphological Productivity*, Cambridge.
- Beito, Olav T. 1954: *Genusskifte i Nynorsk*, Oslo.
- Berntsen, Mandius and Amund B. Larsen 1925/1978: *Stavanger Bymål*, Oslo.
- Blevins, James P. 2015: Inflectional paradigms, in: Matthew Baerman (ed.): *Oxford Handbook of Inflection*, Oxford, p. 87–111.
- 2016: *Word and Paradigm Morphology*, Oxford.
- Boas, Hans 2008: Determining the structure of lexical entries and grammatical constructions in Construction Grammar, in: *Annual Review of Cognitive Linguistics* 6, p. 113–144.
- Booij, Geert 2010: *Construction Morphology*, Oxford.
- Brugmann, Karl 1885: *Zum heutigen Stand der Sprachwissenschaft*, Straßburg.
- Bybee, Joan L. and Carol Lynn Moder 1983: Morphological classes as natural categories, in: *Language* 59, p. 251–270.
- Cameron-Faulkner, Thea and Andrew Carstairs-McCarthy 2000: Stem alternants as morphological signata: evidence from blur avoidance in Polish nouns, in: *Natural Language and Linguistic Theory* 18, p. 813–835.
- Carstairs-McCarthy, Andrew 1994: Inflection Classes, Gender, and the Principle of Contrast, in: *Language* 70, p. 737–788.
- 2008: System-congruity and violable constraints in German weak declension, in: *Natural Language and Linguistic Theory* 26, p. 775–793.
- 2010: *The Evolution of Morphology*, Oxford.
- Conzett, Philipp 2007: *Genus i Norrønt og Nynorsk*. MA thesis, Univ. of Tromsø.
- Corbett, Greville G. 1991: *Gender*, Cambridge.
- Croft, William and D. Allan Cruse 2004: *Cognitive Linguistics*, Cambridge.
- Dąbrowska, Ewa 2004: *Language, Mind and Brain*, Edinburgh.
- 2015: What exactly is Universal Grammar, and has anyone seen it?, in: *Frontiers in Psychology* 6, article 852.
- Dalen, Arnold 1990: Dei trønderske dialektane, in: Ernst Håkon Jahr (ed.): *Den Store Dialektboka*, Oslo, p. 119–141.
- Dammel, Antje 2011: *Konjugationsklassenwechsel*. Berlin.
- Doleschal, Ursula 2000: Gender assignment revisited, in: Unterbeck, Barbara [et al.] (eds.): *Gender in Grammar and Cognition*, Berlin, p. 117–165.
- Enger, Hans-Olav 2004: On the relation between gender and declension: A diachronic perspective from Norwegian, in: *Studies in Language* 28, p. 51–82.
- 2007a: Grammaticalisation due to homonymy avoidance? in: *Transactions of the Philological Society* 105, p. 42–65.
- 2007b: The No Blur Principle meets Norwegian dialects, in: *Studia Linguistica* 61, p. 278–309.
- 2014: Reinforcement in inflection classes: Two cues may be better than one, in: *Word Structure* 7, p. 153–181.

- Fedden, Sebastian and Greville G. Corbett 2020: The German gender system. Presentation at the MultiGender workshop, Oslo, March 2020.
- Fedden, Sebastian [et al.] 2022: Typology meets data-mining: the German gender system. Submitted.
- Fertig, David 2013: Analogy and morphological change, Edinburgh.
- Gaeta, Livio 2018: Natural Morphology, in: Jenny Audring and Francesca Masini (eds.): *The Oxford Handbook of Morphological Theory*, Oxford, p. 244–265.
- Goldberg, Adele. 2006: Constructions at work: the nature of generalization in language, Oxford.
- Golumbia, David 2010: Minimalism is Functionalism, in: *Language Sciences* 32, p. 28–42.
- Gries, Stefan Th. 2018. Mechanistic approaches to language acquisition, in: *Linguistic Approaches to Bilingualism* 8, p. 733–737.
- Haspelmath, Martin 1995: The growth of affixes in morphological reanalysis, in: Geert Booij and Jaap van Marle (eds.): *Yearbook of Morphology 1994*, Dordrecht, p. 1–29.
- Haugen, Odd Einar 1995: *Grunnbok i Norrønt Språk*, Oslo.
- Jackendoff, Ray and Jenny Audring 2018: Relational morphology in parallel architecture, in: Jenny Audring and Francesca Masina (eds.): *The Oxford Handbook of Morphological Theory*, Oxford, p. 390–409.
- Janda, Laura 1996: *Back from the Brink*, München.
- Joseph, Brian 2011: A localistic approach to universals and variation, in: Peter Siemund (ed.): *Linguistic Universals and Language Variation*, Berlin, p. 404–424.
- Kapatsinski, Vsevolod 2018: On the intolerance of the Tolerance Principle, in: *Linguistic Approaches to Bilingualism* 8, p. 738–742.
- Kenstowicz, Michael 1994: *Phonology in Generative Grammar*, Oxford.
- Kiefer, Ferenc 1970: *Swedish Morphology*, Stockholm.
- Köpcke, Klaus-Michael and Verena Wecker 2017: Source- and product-oriented strategies in L2 acquisition of plural marking in German, in: *Morphology* 27, p. 77–103.
- Krahe, Hans and Wolfgang Meid 1969: *Germanische Sprachwissenschaft II: Formenlehre*, Berlin.
- Kürschner, Sebastian and Damaris Nübling 2011: The interaction of gender and declension in Germanic languages, in: *Folia Linguistica* 45, p. 355–388.
- Lasnik, Howard and Terje Lohndal 2010: Government-binding/principles and parameters theory, in: *WIRES Cognitive Science* 1, p. 40–50.
- Lykke, Alexander K. 2019: Variation and Change in the Tense Morphology of Heritage Norwegian in North-America. Ph.D. thesis, University of Oslo.
- Maiden, Martin 1996: The Romance gerund and ›system-dependent naturalness‹ in morphology, in: *Transactions of the Philological Society* 94, p. 167–201.
- 2001: What sort of thing is a derivational affix?, in: Geert E. Booij and Jaap van Marle (eds.): *Yearbook of Morphology 1999*, Dordrecht, p. 25–52.
- 2013: The Latin ›third stem‹ and its Romance descendants, in: *Diachronica* 30, p. 492–530.
- 2018: *The Romance Verb. Morphonic Structure and Diachrony*, Oxford.
- Marcus, Gary F. [et al.] 1995: German inflection: the exception that proves the rule, in: *Cognitive Psychology* 29, p. 189–256.
- Müller, Natascha 2000: Gender and number in acquisition, in: Barbara Unterbeck [et al.] (eds.): *Gender in Grammar and Cognition*, Berlin, p. 351–399.
- Noreen, Adolf 1923: *Altisländische Grammatik I. Halle/Saale*.
- Nübling, Damaris 2000: Prinzipien der Irregularisierung: eine kontrastive Analyse von zehn Verben in zehn germanischen Sprachen, Tübingen.
- Popper, Karl 1972: *The Logic of Scientific Discoveries*, London.

- Tomasello, Michael 2004: What kind of evidence could refute the UG hypothesis? Commentary on Wunderlich, in: *Studies in Language* 28, p. 642–645.
- Venås, Kjell 1967: *Sterke Verb i Norske Målføre*, Oslo.
– 1974: *Linne Verb i Norske Målføre*, Oslo.
- Westergaard, Marit 2009: Usage-based versus rule-based learning: the acquisition of word order in wh-questions in English and Norwegian, in: *Journal of Child Language* 36, p. 1023–1051.
– 2014: Linguistic variation and micro-cues in first language acquisition, in: *Linguistic Variation* 14, p. 26–45.
- Wetås, Åse 2012: Danste, fiskte og hoppte: om apikalsuffiks i preteritum av svake verb i (sørvest-) norske dialektar, in: Unn Røyneland and Hans-Olav Enger (eds.): *Fra Holtijar til Holting*, Oslo, p. 413–424.
- Wittenberg, Eva and Ray Jackendoff 2018: Formalist modeling and psychological reality, in: *Linguistic Approaches to Bilingualism* 8, p. 787–791.
- Wurzel, Wolfgang U. 1984: *Flexionsmorphologie und Natürlichkeit*, Berlin.
– 2000: Morphological properties in the lexicon: diachronic evidence, in: *Acta Linguistica Hungarica* 47, p. 345–356.
- Yang, Charles 2016: *The Price of Linguistic Productivity*, Cambridge.
– 2018: A User’s Guide to the Tolerance Principle, *LingBuzz*.