Journal of Child Language (2021), **48**, 1209–1234
doi:10.1017/S0305000920000732

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Productivity and the acquisition of gender

Sigríður Mjöll BJÖRNSDÓTTIR

The Arctic University of Norway, Tromsø
Address for correspondence: sigridur.m.bjornsdottir@uit.no

(Received 14 December 2019; revised 14 July 2020; accepted 9 November 2020;
first published online 4 February 2021)

Abstract

Children's differing learning trajectories cross-linguistically have been at the forefront of gender acquisition research, often with conflicting results and conclusions. As a result, the source of children's different learning behaviors in gender acquisition has been unclear. I argue that children's gender acquisition is driven by the search for productive patterns. First, I provide corpus studies where the predictions of a learning model (Yang, 2016) are formulated. Second, I report the results of an elicited production task on Icelandic-speaking children (N = 26, ages 2;6-6;3 years) and adults (N = 18) that puts these predictions to test. The results suggest that Icelandic-speaking children and adults draw a categorical distinction between productive and unproductive suffixes in Icelandic gender assignment. I discuss the implications of these findings for morphological learning beyond gender acquisition.

Keywords: gender; morphology; rules; defaults; productivity

Introduction

Grammatical gender has conventionally been defined as the sorting of nouns into classes as reflected in agreement morphology (Corbett, 1991; Hockett, 1958). Gender systems differ cross-linguistically with respect to what kind of information is predictive of gender assignment. A distinction has been made between STRICT SEMANTIC SYSTEMS, as exemplified by the gender systems of the Dravidian languages, and FORMAL SYSTEMS, as exemplified by typologically diverse languages, such as Qafar and Russian (Corbett, 2013). Given the typological diversity of gender systems, children must be able to detect a wide range of formal and semantic regularities on the basis of language-specific data.

In her seminal study, Karmiloff-Smith (1979) showed that French children were able to assign gender on the basis of noun endings. Moreover, the children seemed to rely on noun endings even if the resulting gender were at odds with the biological sex of the referent. Similar results have been obtained many times cross-linguistically (Clark, 1985; Hernández-Pina, 1984; Levy, 1983; Mills, 1986; Rodina & Westergaard, 2012; 2013; 2015). Collectively, the results of this body of research suggest that children

can learn gender systems that are detached from any semantic motivation. However, research on more typologically diverse gender systems is needed in order to determine whether this early formal bias is an artifact of the language sample or a finding about early grammatical representation.

Children's learning trajectories of grammatical gender vary cross-linguistically (Mills, 1986). Gender systems have been divided into two groups from an acquisitional perspective: Transparent and opaque (Slobin, 1977). Transparent gender systems have a set of productive patterns for gender assignment, whereas opaque gender systems have few or none. Productive rules in transparent systems, such as Spanish and Russian, are typically in place by the age of three (Lew-Williams & Fernald, 2007; Rodina & Westergaard, 2012), whereas the paucity of such rules in opaque systems, like Norwegian and Dutch, results in late mastery (Rodina & Westergaard, 2013; Unsworth & Hulk, 2010). Transparent or opaque, gender acquisition involves detecting language-specific patterns and evaluating whether they are useful for learning or not. In other words, the child learner must somehow outweigh the evidence for and against a pattern in order to determine whether or not it can be used to form a generalization about gender assignment.

Even within a transparent gender system, gender assignment rules[1] may be learned at different rates. Mills (1986) proposed, using evidence from German, that gender assignment rules were acquired in order of CLARITY. By her definition, clarity is determined by the scope of the rule and the number of exceptions; the greater the scope of the rule and the fewer exceptions, the earlier the rule is acquired. For example, she argued that the rule with the greatest scope in German is "nouns that end in –e are feminine" because of the high frequency of the pattern and the low number of exceptions (p. 85). However, even the role of frequency has been debated. For instance, Henzl (1975) argued, using evidence from Czech, that children first formulated gender assignment rules on the basis of noun endings which are "least ambiguous", irrespective of frequency.

Hitherto it has been unclear what makes a gender system either transparent or opaque to the child learner. In parallel, it has been unclear how the child learner can determine the scope of a gender assignment pattern. Therefore, a theory of gender acquisition is needed that can both identify the conditions under which a gender assignment pattern is useful to the learner – and when these conditions are not met.

In this paper, I propose an approach whereby gender acquisition is characterized by a search for productive gender assignment rules guided by a learning model (Yang, 2005; 2016). First, I discuss prior research on productivity in first language acquisition. Second, I introduce the Tolerance Principle, a quantitative model of productivity (Yang, 2005; 2016). I discuss the relevance of quantitative methods for research on gender acquisition and demonstrate how the approach works using grammatical gender in Spanish as a test case. Next, I show how predictions for Icelandic gender acquisition can be made on the basis of child-directed speech and child naturalistic data. Moreover, I show how these predictions robustly hold when samples are created from other corpora to approximate children's vocabulary size during the stages of gender acquisition. Subsequently, I present the results of an elicited production task on Icelandic children and adults. Finally, I discuss an alternative view of productivity (Baayen, 1989; 1992; 1993) and evaluate its predictions against the empirical results. The paper concludes with a discussion of the implications of these findings for morphological learning beyond gender acquisition.

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

Productivity and absence thereof in language acquisition

Language acquisition involves learning words and how to inflect them. The source of children's ability to learn inflectional patterns has been a point of contention for theories of morphological learning. In her famous Wug experiments, Berko (1958) showed convincingly that English-speaking children extend productive inflectional patterns like, for example, the plural suffix –s, when inflecting novel words. Children have also been found to over-generalize productive patterns in naturalistic settings even though this may result in forms that are not attested in the input, such as *foots and *breaked (Pinker & Prince, 1994). Children's ability to extend productive patterns in both experimental and naturalistic settings has been taken as evidence for rule-based learning in acquisition.

However, sometimes productivity fails. Gaps within an inflectional paradigm are the result of having no acceptable morphological option or default (Baronian & Kulinich, 2012; Halle, 1973; Fanselow & Féry, 2002; Orgun & Sprouse, 1999; Pertsova, 2005). Morphological gaps are common cross-linguistically. For instance, many English speakers find the past participles of certain irregular verbs, like stride, problematic (Pinker, 1999). Similarly, there are no acceptable 1SG forms for a handful of verbs in Spanish (Albright, 2003). There are no semantic reasons for this ineffability. Rather, it seems to reflect speakers' failure to generate a systematic pattern or a rule. Morphological gaps have posed a challenge to rule-based accounts, as the unavailability of a rule or a default form is unexpected.

The learning trajectory of Polish noun inflection suggests that children do not need to resort to defaults in order to learn inflectional morphology (Dabrowska, 2001; 2005). Polish nouns are inflected for gender, case and number. The most important factor in determining the choice of inflectional ending is gender (Dabrowska, 2001, p. 558). The most interesting case is the choice of ending for masculine genitive singular nouns: masculine singular nouns in Polish can take either –a or –u as a genitive ending in a seemingly unpredictable fashion. While –a is the most frequent masculine genitive singular ending, it does not seem to have the status of a default, since loanwords and low frequency masculine singulars can take either ending.

In a series of longitudinal corpus case studies, Dabrowska (2001) showed that Polish noun inflection was largely in place by the age of 2;0. Furthermore, Polish-speaking children made few errors with masculine genitive singular nouns in spite of the arbitrary distribution of the two endings. In case of errors, children made unsystematic choices of either ending.

These findings have been taken as evidence against rule-based learning (Clahsen, 1999; Pinker, 1999). Instead, Dabrowska (2001, 2005) argued that they lent support to USAGE-BASED approaches to language acquisition (Tomasello, 1992; 2003). Hence, the absence of productivity has raised key questions about the nature of the mechanism underlying linguistic creativity.

Predicting productivity and absence thereof

The Tolerance Principle

There is general agreement that language has both productive and unproductive patterns. However, the division line between the two has been a point of contention. Yang (2005; 2016) has proposed a model of linguistic productivity, the Tolerance

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

Principle, to account for how children distinguish between productive and unproductive patterns on the basis of positive evidence in the input. The Tolerance Principle quantifies the precise conditions for productive rule formation. The model hypothesizes that a general rule will be formed when doing so is computationally more efficient than storing lexical forms. The principle is stated in (1).

(1) The Tolerance Principle
If R is a productive rule applicable to N candidates, then the following relation holds between N and e, the number of exceptions that could but do not follow R:

$$e \leq \theta N \text{ where } \theta N = N/\ln N$$

The Tolerance Principle states that it is computationally more efficient to form a productive rule only when the number of exceptions is less than the number of items divided by the natural log of the number of items. Computational efficiency is computed by calculating the time complexity required for forming a rule with the time complexity required for accessing individual lexical forms. Crucially, the division between productive and unproductive processes is a categorical one on this approach.

The Tolerance Principle makes use of the Elsewhere Condition (Kiparsky, 1973), which states that when a more specific form (or rule) is available, it is preferred over a more general one. For example, went is the past tense form for the verb go, so it overrides the regular but ungrammatical *goed. The Elsewhere Condition is implemented by the Tolerance Principle as a serial search procedure, which is empirically motivated by research on language processing (see Yang, 2016, pp. 49–60).

To illustrate this serial procedure, one can think of past tense acquisition in English. The child is faced with verbs that adhere to the regular pattern, "add -d", and verbs that do not. The Tolerance Principle assumes that, in order to be maximally efficient in forming the past tense of verbs in English, the child is faced with two options: 1) Store all past tense verb forms individually 2) Form a productive rule. In the first case scenario, every item is stored in a list ranked by frequency. This means that the learner must search the list every time there is an occasion to express the past tense of a verb. In the second case scenario, only the exceptions are stored in a frequency-ranked list. The list of exceptions must be searched first before the productive rule can be applied.

The Tolerance Principle operates on type counts. Therefore, productivity in grammar learning on this approach is connected to the number of types over which linguistic patterns are expressed, rather than the number of tokens. The same view has been adopted by a wide variety of research programs (Aronoff, 1976; Baayen, 1993; Bybee, 1985; Plunkett & Marchman, 1991).

Given a well-defined hypothesis space, the Tolerance Principle can be used as a quantitative measure to predict whether any given linguistic pattern can be perceived by the child learner as productive or not. The Tolerance Principle is just one thresholding function and has a wide range of empirical support (consult Yang, 2016 for case studies). In addition, the predictions of the Tolerance Principle have been borne out for children in experimental settings (Schuler et al., 2016).

Language acquisition involves not only detecting productive patterns, but also unproductive patterns. The Tolerance Principle not only models the conditions for productive rule formation; it can also identify conditions under which no

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

productive rule is available (Gorman & Yang, 2018). For example, the Tolerance Principle can predict the absence of a default genitive ending for Polish masculine singulars on a numerical basis. Table 1 shows the numerical distribution of Polish masculine genitive singular nouns by ending (adapted from Yang, 2016, based on CHILDES).

Table 1. Numerical Distribution of Genitive Endings for Masculine Singular Nouns in Polish

	Suffix	Types	Productive?
GEN.SG	*–a*	837 (62%)	No (516>188)
	–u	516 (38%)	No (837>188)
Total		1353	
θ_N		1353/ln1353=188	

An analysis using the Tolerance Principle revealed that in spite of the statistical majority of –a as the genitive ending of masculine singulars, the number of nouns that take the alternative ending is too great for –a to be productive. On this approach, therefore, absence of productivity does not constitute as evidence against rule-based learning. Rather, it is the direct consequence of a learning process guided by a search for productivity that fails to succeed and results in rote memorization.

Relevance to gender acquisition

Approaches using quantificational methods have the advantage of being able to make clear, testable predictions on the basis of input data. In this section, I will briefly showcase how the present approach works using the Spanish gender system as an example.

The Spanish gender system distinguishes between masculine and feminine nouns. There are correlations between nominal morphology and gender assignment: Nouns that take the suffix –o tend to be masculine, whereas nouns that take the suffix –a tend to be feminine. In an eye-tracking study, Lew-Williams and Fernald (2007) showed that Spanish-learning children, aged 2;10–3;6 years, were able to use gender-marked articles to establish reference of such nouns. Thus, young Spanish-learning children had internalized productive gender assignment rules in spite of an estimated vocabulary of only 500 words.

The distribution of noun types across gender and suffix in a longitudinal corpus of Spanish child-directed speech (Linaza et al., 1981) is provided below in Table 2. The corpus reflects the interaction between a caregiver and their child between the ages of two and four. Therefore, it should give a reasonable estimate of a child's vocabulary size in Spanish gender acquisition.

An analysis using the Tolerance Principle confirmed the productivity of –o to masculine and –a to feminine. In the absence of a suffix, the Tolerance Principle predicted masculine to be the default gender in Spanish.

These predictions are consistent with studies on Spanish gender acquisition in both naturalistic and experimental settings: Children generalize masculine to nouns with the suffix –o and feminine to nouns with the suffix –a. In the absence of a productive suffix,

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

Table 2. Numerical Distribution of Noun Types by Gender and Suffix in Spanish Child-Directed Speech

Suffix	M	F	θ_N	Productive?
–o	113	2	115/ln115=24	Yes (24>2)
–a	8	116	124/ln124=26	Yes (26>8)
–ø	102	16	118/ln118=25	Yes (25>16)
Total	223	134		

they resort to the default gender: namely, masculine (see, among many, Clark, 1985; Hernández-Pina, 1984; Mariscal, 2008; Pérez-Pereira, 1991).

The Icelandic gender system

Icelandic has a gender system that distinguishes between masculine, feminine and neuter. Typologically, the Icelandic gender system has been classified as formal (Corbett, 2013). Icelandic has rich agreement morphology that manifests itself on the definite article, which is a suffix (2a), adjectives (2b), the past participle (2c) and pronouns (2d). Anaphoric pronouns must refer to the formal gender of the referent noun irrespective of animacy or biological sex.

(2) a. Stóll-**inn**, skál-**in**, borð-**ið**.
Chair-M.DEF, bowl-F.DEF table-N.DEF
'The chair, the bowl, the table.'

b. Flott-**ur** stóll, flott-**ø** skál, flott-**ø** borð.
Nice-M chair-M, Nice-F bowl-F, nice-N table-N
'A nice chair, a nice bowl, a nice table.'

c. Stóllinn er brot-**inn**, skálin er brot-**in**,
The chair-M is broken-M, the bowl-F is broken-F,
borðið er brot-**ið**.
the table-N is broken-N
'The chair is broken, the bowl is broken, the table is broken.'

d. **Hann** er brotinn, **hún** er brotin, **það** er brotið.
He is broken, she is broken, it is broken.
'He (the chair) is broken, she (the bowl) is broken, it (the table) is broken.'

The three genders are roughly equally frequent: 32% are masculine, 38% feminine and 30% are neuter (Helgadóttir et al., 2010). These numbers are consistent with the input corpora that will be examined later in the paper.

In addition to gender, Icelandic distinguishes between four cases: Nominative, accusative, dative and genitive. Gender and inflection in Icelandic interact to form INFLECTION CLASSES, which are standardly defined as a set of roots that each share the same set of inflectional realizations (Aronoff, 1994).

Icelandic reference grammars (see e.g., Kvaran, 2005) have standardly followed the lead of Old Norse reference grammars (Iversen, 1922; Noreen, 1903) by stating the correspondence between gender and inflection without discussing specific gender

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

assignment rules. The idea is that the gender of a noun can be determined by its inflection class membership to some extent.

Nominative singular is the most frequent inflectional form in Icelandic, constituting 40% of all nominal forms (Helgadóttir et al., 2010). Furthermore, due to syncretism in the nominal paradigm, many forms are identical to the nominative singular in oblique cases. There are strong correlations between nominative singular morphology and gender assignment in Icelandic as in other fusional languages like, for example, German and Russian (Corbett, 1991). In particular, three nominative singular suffixes are predictive of either masculine or feminine, respectively.[2]

(3) a. Nouns that take the nominative singular suffix –r are typically masculine.[3]
b. Nouns that take the nominative singular suffix –i are typically masculine.
c. Nouns that take the nominative singular suffix –a are typically feminine.

Table 3 demonstrates how these suffixes map on to real nouns in Icelandic.

While these patterns are robust in Icelandic, they do have exceptions. For instance, some feminine nouns take the nominative singular suffix –r. Diachronically, most of these nouns have shifted to masculine (Iversen, 1922; Noreen, 1903).

Table 3. Mappings between Gender and Nominative Singular Suffixes in Icelandic

	–r	*–i*	*–a*
Masculine	sokku-r (‘a sock’)	burst-i (‘a brush’)	NA
Feminine	brúðu-r (‘a bride’)	NA	fat-a (‘a bucket’)
Neuter	NA	NA	NA

The absence of an overt nominative singular suffix is indicated by -ø. Some nouns do not take the phonemes in Table 3 by suffixation. Instead, they form part of the noun‘s stem, as shown in (4). These nouns tend to have low type but high token frequency. Most of these nouns are neuter, although nouns with stem-final /i/ can be either feminine or neuter (4b).

(4) a. Aug**a**-ø, eyr**a**-ø.
Eye-N.NOM.SG, ear-N.NOM.SG
‘An eye, an ear.’

b. Tæk**i**-ø, gleð**i**-ø.
Device-N.NOM.SG, joy-f.NOM.SG
‘A device, joy.’

c. Be**r**-ø, ke**r**-ø.
Berry-N.NOM.SG, tub-N.NOM.SG
‘A berry, a tub.’

While these nouns have oblique forms different from nouns that take these sounds by suffixation, they could be ambiguous to the child learner in gender acquisition given the statistical dominance of nominative singular forms in the input. Therefore, these nouns are counted as exceptions to the general patterns stated in (3) in subsequent quantitative analyses.

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

The choice of nominative singular suffix is a result of morphological, rather than phonological selection. The same root may select for more than one suffix to yield a minimal pair as in (5a). Some borrowed nouns show variation in the choice of suffix, which in turn affects gender assignment (cf. 5b-c).

(5) a. Sæt-**i**, sæt-**a**.
Cutie-M, cutie-F
'Male cutie, female cutie.'

b. Djóku-**r**, Djók-**ø**.
Joke-M, joke-N
'A joke.'

c. lúpp-**a**, lúpp-**ø**.
loop-F, loop-N
'A loop.'

There is no productive nominative singular suffix for neuter nouns. The stem-final segment of neuter nouns can consist of any phonotactically legal consonant or a vowel (see above). There are no clear phonological patterns specific to neuter. For instance, many neuter monosyllabic nouns rhyme with feminine monosyllabic nouns.

(6) a. Borg-ø, torg-ø.
city-F, square-N
'A city, a square.'

b. Ull-ø, gull-ø,
wool-F, gold-N
'Wool, gold.'

Neuter has standardly been assumed to be the default gender in Icelandic (Steinmetz, 1985). This assumption will be challenged later in this paper.[4]

Most nouns in Icelandic are assigned only one gender. In case of variation in gender assignment, however, nouns that lack an overt nominative singular suffix are the primary targets. These nouns have also undergone gender shifts diachronically (Noreen, 1903; Iversen, 1922). The attested variation seems arbitrary. Similarly, there is both inter-speaker and intra-speaker variation in the gender assignment of some borrowed nouns in Icelandic. Thus, while the choice of nominative singular suffix clearly determines the gender of both jeppi and paranója, the absence of such a suffix seems to correlate with variation in gender assignment, as shown in Table 4.

Table 4. Gender Assignment of Borrowed Nouns in Icelandic

Noun	Gender	Meaning
Jepp-i	M	'A jeep'
Paranój-a	F	'Paranoia'
Jógúrt-ø	M, F, N	'Yoghurt'
E-mail-ø	M, N	'E-mail'

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

To conclude this section; given the statistical dominance of nominative singular morphology, it seems plausible to assume that Icelandic children learn these inflectional patterns early and use them as base forms in gender acquisition.

Gender acquisition in icelandic: a longitudinal corpus case study

Data

The data consist of longitudinal recordings of a caregiver's speech to an Icelandic-speaking child and the child's spontaneous speech in response (Sigurjónsdóttir, 2007). A total of 82 recordings were made approximately once a month when the child was between the ages of 1;6–4;3 years. The child-directed speech contained around half a million tokens; whereas the child's spontaneous speech contained around 7000 tokens.

Procedure

Nominative singular noun types were extracted from the corpus and tagged for gender and suffix. Child and adult data were analyzed separately. The purpose of the child analysis was to test whether the same predictions could be made on the basis of the child's vocabulary. Both child and adult data were subjected to a quantitative analysis using the Tolerance Principle. In addition, the child naturalistic data was subjected to an error analysis.

Analysis of child-directed speech

The caregiver's speech contained 478 nominative singular noun types, which constituted approximately 41% of all noun types that were produced. Their numerical distribution by gender and suffix is provided in Table 5. Token numbers are given in brackets.

Table 5. Numerical Distribution of Nominative Singular Noun Types in Icelandic Child- Directed Speech

NOM.SG	M	F	N	θ_N	Productive?
–R	63 (494)	3 (53)	4 (57)	70/ln70=16	Yes (16>7)
–i	82 (449)	4 (55)	8 (218)	94/ln94=21	Yes (21>12)
–a	0	133 (593)	4 (10)	137/ln137=28	Yes (28>4)
–ø	29 (144)	35 (507)	134 (721)	198/ln198=37	No (37<64)
Total	174	175	150		

Both nominative singular suffixes –r and –i were predicted to be productive of masculine by the Tolerance Principle, as the number of non-masculine nouns with these suffixes was below the exception threshold (θN). Likewise, –a was predicted to be productive of feminine.

In the absence of a nominative singular suffix, however, no gender was predicted to be productive. Thus, in spite of the statistical dominance of neuter within this category,

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

the number of non-neuter nouns exceeded the exception threshold. As a result, Icelandic was predicted to lack a default gender in the absence of a productive nominative singular suffix.

Analysis of child naturalistic production

The child produced a total of 345 nominative singular noun types, which constituted approximately half of all noun types that were produced. Their numerical distribution by gender and suffix is provided in Table 6. Token numbers are given in brackets.

Table 6. Numerical Distribution of Nominative Singular Noun Types in Child Naturalistic Production

NOM.SG	M	F	N	θ_N	Productive?
–R	47 (167)	3 (4)	2 (12)	52/ln52=13	Yes (13>5)
–i	41 (143)	2 (11)	4 (33)	47/ln47=12	Yes (12>6)
–a	0	97 (221)	4 (18)	101/ln101=22	Yes (22>4)
–ø	30 (108)	36 (122)	55 (178)	121/ln121=25	No (55<66)
Total	118	138	65		

The same predictions were made on the basis of the child's spontaneous speech as on the child-directed speech, even if the child's production contained fewer noun types. The child was predicted to have internalized three productive rules of gender assignment in the absence of a default gender.

Error analysis of child naturalistic speech

The child was 100% target-consistent with nouns that take suffixes –r, –i and –a in the corpus. This means that the child had internalized the gender of these nouns before their second birthday. The child's non-target-consistent gender agreement exclusively targeted nouns that had no overt nominative singular suffix (–ø), with an error rate of 4.6%. The nouns affected alternated between all three genders. Examples of this are provided below in Table 7.

Table 7 Non-Target-Consistent Gender Agreement in Icelandic Child Naturalistic Production Child Production

Child Production	Target-Consistent Use	Meaning
*Krús-ið mitt mug-N.DEF my-N	Krús-in mín mug-F.DEF my-F	'My mug'
*Falleg-ø kjóll beautiful-F dress-F	Falleg-ur kjóll beautiful-M dress-M	'A beautiful dress'
*Úr-inn minn watch-M.DEF my-M	Úr-ið mitt watch-N.DEF my-N	'My watch'

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

The child's non-target consistent gender agreement did not suggest the application of a default gender. Rather, the pattern attested appeared unsystematic.

Corpora as an estimate of linguistic experience

Corpus data is a sample of linguistic experience. Any two sets of corpora are unlikely to contain the exact same linguistic items. This is analogous to child language acquisition; children's linguistic experience is inevitably variable.

So far, the corpus analyses in this paper have been based on small corpora. However, a small vocabulary is developmentally appropriate in the study of gender acquisition. Gender, in languages with productive gender assignment rules, is largely in place by the age of three when children typically know only a few hundred words (Hart & Risley, 1995; 2003; Szagun et al., 2006). The question is how children can converge on the target gender system on the basis of a vocabulary that is both small and variable from child to child.

One way to address this question is to study differences between corpora of different sizes and genres. Kodner (2019) studied the differences between corpora derived from adult literary genres and child-directed speech in a series of case studies. He found that once adult literary corpora had been trimmed by frequency, they had statistically similar type counts to child-directed speech corpora in spite of lexical differences. In other words, the main difference between adult literary corpora and child-directed speech involved low frequency lexical items. One implication of these findings is that children's grammar learning may be based on high frequency lexical items, rather than adult-size lexicons.

In this section, predictions will be made using the Tolerance Principle on the basis of an adult online corpus. The objective is to establish whether the same predictions can be made when lexical items are drawn at random using a computer simulation model from a much larger language sample.

Furthermore, predictions will be formulated on the basis of the top few hundred most frequent noun types.

Data

The data consist of a corpus of 8.6 million tokens (https://github.com/hermitdave/FrequencyWords/blob/master/content/2018/is/is_full.txt) that were extracted from the SUBTLEX corpus (http://www.opensubtitles.org/). Corpora based on subtitles have been shown to be a good approximation of spoken languages (https://www.ugent.be/pp/experimentele-psychologie/en/research/documents/subtlexus).

Procedure

A computer simulation model was run on the corpus. The model was instructed to draw 500,000 noun tokens, to match the token size of the Icelandic child-directed speech corpus, at random and proportionally to word frequencies. Noun types that occurred less frequently than once per million words were excluded from the analysis. Nominative singular noun types were extracted from the sample and categorized by gender and suffix. They were then subjected to a quantitative analysis using the Tolerance Principle.

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

Results

563 nominative singular noun types were attested in a random sample of 500,000 words in the SUBTLEX corpus. Their numerical distribution by gender and suffix is provided in Table 8. Token numbers are given in brackets.

The Tolerance Principle made the same predictions based on the SUBTLEX corpus as on Icelandic child-directed speech (cf. Table 5) in spite of differences both in terms of lexical items and type counts.

Table 8. Distribution of Noun Types by Gender and Suffix in the SUBTLEX Corpus

NOM.SG	M	F	N	θ_N	Productive?
–r	134 (1483)	4 (66)	5 (25)	143/ln143=29	Yes (29>9)
–i	97 (746)	4 (62)	17 (534)	118/ln118=25	Yes (25>17)
–a	1 (617)	92 (808)	2 (15)	95/ln95=21	Yes (21>3)
–ø	20 (596)	69 (1392)	125 (1441)	214/ln214=40	No (40<89)
Total	252	169	142		

Formulating predictions for small vocabularies

Table 9 shows the predictions of the Tolerance Principle on the basis of the top 100 and top 300 most frequent nominative singular noun types in the SUBTLEX corpus.

Table 9. Distribution of the most Frequent Noun Types in the SUBTLEX Corpus by Gender and Suffix

	M		F		N		θ_N		Productive?	
NOM.SG	Top 100	Top 300	Top 100	Top 300	Top 100	Top 300	Top 100	Top 300	Top 100	Top 300
–r	20	68	2	5	0	3	7	17	Yes	Yes
–i	24	47	1	7	2	9	8	14	Yes	Yes
–a	0	1	22	74	1	5	7	18	Yes	Yes
–ø	10	15	9	27	14	28	9	16	No	No

The Tolerance Principle made the same predictions as before, irrespective whether the analysis was based on the top 100 or top 300 most frequent noun types.

Discussion

Children's linguistic experience is inevitably variable: Children are unlikely to know the exact same words and their vocabulary sizes differ, even for children at the exact same age. In spite of lexical differences, however, children acquiring the same language are able to discover what the target grammar is.

The Tolerance Principle operates on types. As a consequence, what matters for learning is the number of lexical items that exhibit a specific property, rather than which exact lexical items those are. In this section, I have shown that, while the type

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

counts of grammatical properties may differ from corpus to corpus, the predictions are the same. This is because the proportion of exceptions that go against a linguistic pattern relative to the types that conform to a linguistic pattern yields the same results, regardless of the exact number of types involved in the calculations.

Child-directed speech and adult corpora have been shown to converge on high frequency lexical items (Kodner, 2019). Therefore, it is plausible that children base their grammar learning mainly on high frequency lexical items. An analysis of the most frequent noun types in the SUBTLEX corpus using the Tolerance Principle predicted an early division between productive and unproductive suffixes in Icelandic gender assignment.

Experimental study

Participants

26 children (M = 4;5 years, SD = 1.33 years, age range = 2;9–6;3 years; 14 females, 12 males) and eighteen adult controls participated in this study. An additional four children participated, but were excluded from analysis due to failure to understand the task or unwillingness to engage with the game. Children were recruited from a day-care centre in suburban Reykjavík, where the study was conducted. Adult participants were recruited at the University of Iceland, Reykjavík. All participants were native speakers of Icelandic with normal hearing and normal to corrected-to-normal vision. No participant identified as bilingual/multilingual or reported to have a history of language delay.

Design

An elicited production task was designed with two conditions: Productive and Unproductive. In the Productive condition, participants were exposed to a novel noun with either suffix –r, –i or –a. In the Unproductive condition, participants were exposed to a novel noun, monosyllabic or disyllabic, that did not bear such a suffix.

Predictions

The Tolerance principle predicted that participants would make categorical suffix-based choice in gender assignment in the Productive condition, but arbitrary gender choices in the Unproductive condition.

Materials

28 nonce nouns were designed. The novel nouns all conformed to phonetic and phonological restrictions in Icelandic. To control for phonological neighbourhood density, the Phonological Corpus Tools software (Hall et al., 2016) was used to check for minimal pairs with nouns included in Pind's (1991) frequency list of Icelandic. The stem-final segment of novel nouns in the Unproductive condition could be any consonant except /r/. The novel nouns are given in Table 10.

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

Table 10. Test Items by Nominative Singular Suffix

–r	*–i*	*–a*	*–ø* MS	*–ø* DS
Lerfur	Krandi	Bukla	Rúf	Kútes
Sappur	Lerri	Rala	Kurk	Ratef
Mækur	Tukki	Húla	Mirg	Farem
Tirgur	Dubbi	Darga	Súf	Múkaf
		Múka	Glæf	Silám
		Fóma	Lirg	Rútis
			Turk	Tækill

The novel nouns were paired with inanimate novel objects from the Novel Object and Unusual Name (NOUN) database (Horst & Hout, 2016). Figure 1 shows an example of a novel object used in the study:

Figure 1. A Novel Object at Exposure to Test

There were fourteen test items per condition. The test items were organized into seven trials. In each trial, the participant was presented with four test items, two for each condition, in a randomized order.

The test sentence served the purpose of a magical charm to be uttered by the participant in lieu of more traditional charms like 'hocus pocus'. The construction induced gender agreement on the definite suffix and possessive pronominal, as shown for real nouns in (7):

(7) a. Hvar er hattu-r-**inn minn**?
where is hat-M.DEF.SG my-M
'Where is my hat?'

b. Hvar er penn-i-**nn minn**?
where is pen-M.DEF.SG my-M
'Where is my pen?'

c. Hvar er kann-a-**n mín**?
where is mug-F.DEF.SG my-F
'Where is my mug?'

d. Hvar er egg-**ið mitt**?
where is egg-N.DEF.SG my-N
'Where is my egg?'

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

The construction was chosen in light of the fact that children acquiring Icelandic have been shown to comprehend and produce main clause *wh*-questions early. Moreover, *wh*-questions with *where* are among the earliest interrogative questions attested in Icelandic child language, with no reported erroneous use (Sigurjónsdóttir, 1991).

Procedure

The task was embedded in an animated interactive movie that was played off a computer screen. The movie was designed using Animaker, an online animation video maker and was thirteen minutes long. Children and adults were tested individually in a quiet location at a day care center and at the University of Iceland.

The objective of the task was to help the movie's story protagonist obtain novel toys by magic. However, in order for the novel toys to come to be obtained, the participant had to be able to use the name of the novel toy in a sentence at test. The participant was shown a picture of the novel object and heard its name twice in syntactic contexts where the nominative singular is obligatory, as (8) demonstrates.

(8) a. Þetta er **lerfur**.
this is lerfur-M.NOM.SG.
'This is a lerfur.'

b. Vá! **Lerfur**!
wow lerfur-M.NOM.SG.
'Wow! A lerfur!'

After the participant had produced the test sentence, the novel object appeared by magic as shown in Figure 2.

Figure 2. Magic at work in the Test Scene

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

Prior to test, there was a training session in which the participant observed the story protagonist either succeed or fail with the magic. The purpose of these scenes was to provide the participant with both positive and negative reinforcement. Subsequently, the participant was trained on three real nouns of each gender.

Results

Children

Children's behavior across the two conditions is summarized in Figure 3. Dots represent individual performance in each condition. Bars are standard error. Productive gender assignment in the Productive condition corresponds to mean systematic suffix-based choice of gender: Masculine for nouns with *–r* or *–i*, feminine for nouns with *–a*. In order to confirm the unproductivity of neuter in Icelandic, productive gender assignment in the Unproductive condition corresponds to mean neuter assignment.

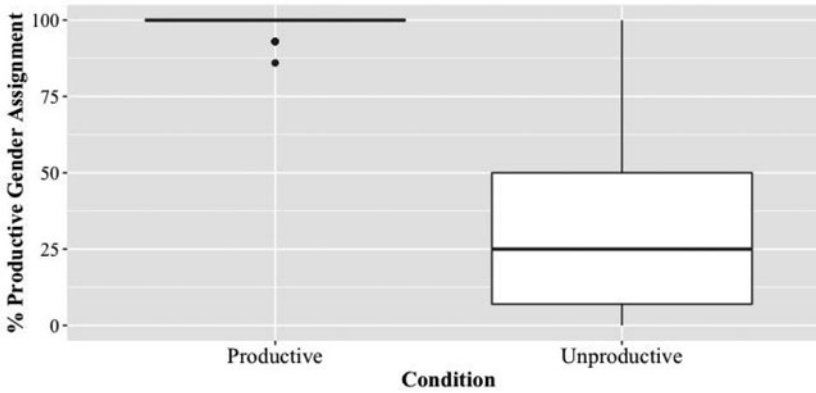

Figure 3. Children: Gender Assignment across Conditions

Children made a categorical, suffix-based choice of either masculine or feminine in the Productive condition. They assigned masculine consistently to novel nouns with either suffix *–r* or *–i* ($M = 0.99$, $SD = .037$, $SE = .007$). Likewise, they assigned feminine consistently to novel nouns with the suffix *–a* ($M = 0.98$, $SD = .04$, $SE = .009$). The percentage of neuter assignment in the Productive condition was 2.35%, which is not statistically significant from zero. In the Unproductive condition, children did not make a systematic choice of neuter ($M = 0.29$, $SD = 0.28$, $SE = .05$). A paired t-test confirmed a significant difference between the means of the two conditions: $t(25) = 11.93$, $p < .001$.

Figure 4 shows the distribution of children's responses in the Unproductive condition. Omission was defined as silence at test. Variable assignment was defined as the repetition of a test item twice, or more often, with variable gender agreement.

Gender assignment in the Unproductive condition was characterized by a great deal of inter-and intra-speaker variation. Collectively, the children did not behave categorically in this condition, although six children did make categorical choices of

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

gender. Nevertheless, these children were categorical in different ways: Three assigned feminine categorically or near-categorically, two assigned masculine categorically and one assigned neuter categorically.

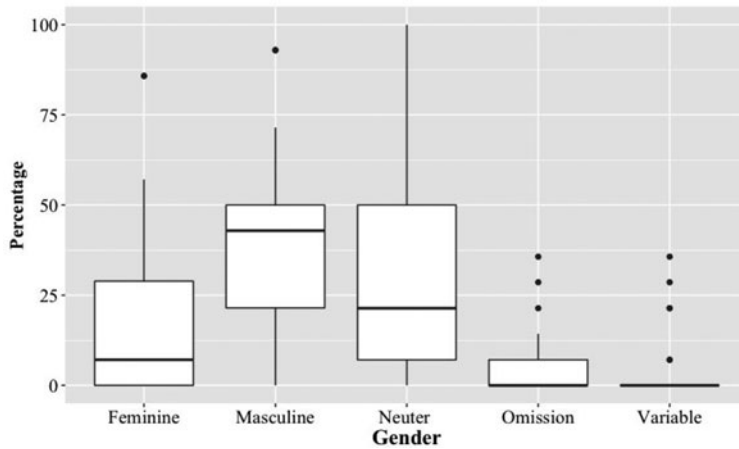

Figure 4. Children: Gender Assignment in the Unproductive Condition

A paired t-test revealed no significant difference between mean neuter assignment of monosyllabic and disyllabic nouns: $t(24) = -0.52$, $p = 0.61$. Figure 5 shows gender assignment of monosyllabic and disyllabic nouns in the Unproductive condition.

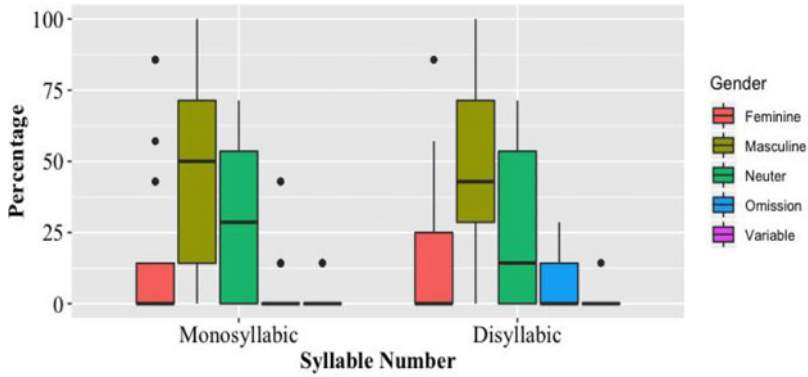

Figure 5. Children: Gender Assignment and Syllable Number in the Unproductive condition

In order to assess the relationship between age and neuter assignment, a simple regression analysis was conducted. The relationship is visualized in Figure 6. The result of the analysis showed no correlation between age and mean neuter assignment ($r = .09$).

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

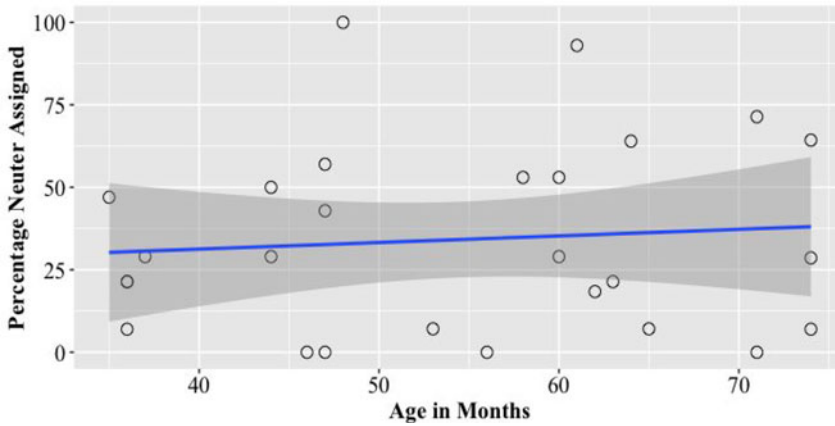

Figure 6. Effect of Age on Neuter Assignment

Adults

Adults' behavior across the two conditions is summarized in Figure 7. Dots represent individual performance in each condition. Bars are standard error. As before, productive gender assignment in the Productive condition corresponds to mean systematic suffix-based choice of gender: Masculine for nouns with *–r* or *–i*, feminine for nouns with *–a*. In order to confirm the unproductivity of neuter in Icelandic, productive gender assignment in the Unproductive condition corresponds to mean neuter assignment.

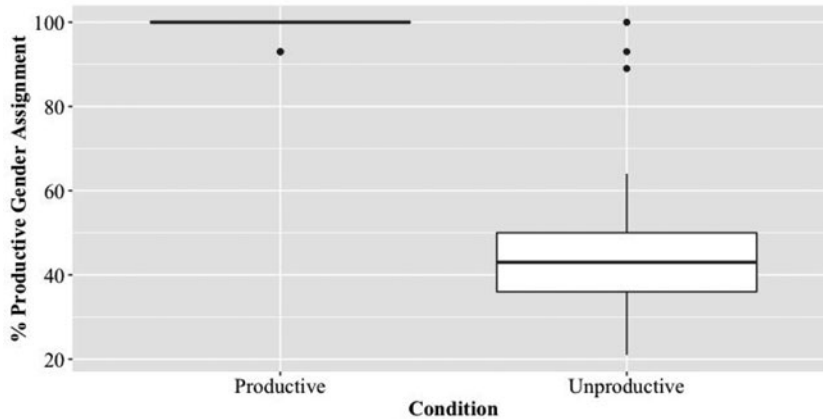

Figure 7. Adults: Gender Assignment across Conditions

Adults made a categorical, suffix-based choice of either masculine or feminine in the Productive condition. They assigned masculine at ceiling (100%) to novel nouns with either suffix *–r* or *–i*. Similarly, they assigned feminine consistently to novel nouns

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

with the suffix *–a* ($M = 0.99$, $SD = .03$, $SE = .009$). Mean neuter assignment in the Unproductive condition was 48% ($SD = 0.24$, $SE = .013$). A paired t-test confirmed a significant difference between the two conditions: $t(17) = 9.32$, $p < .001$.

Figure 8 displays the distribution of adults' responses in the Unproductive condition. Gender assignment in the Unproductive condition was characterized by inter-and intra-speaker variation. Collectively, adults did not behave categorically in this condition, although three chose consistently neuter.

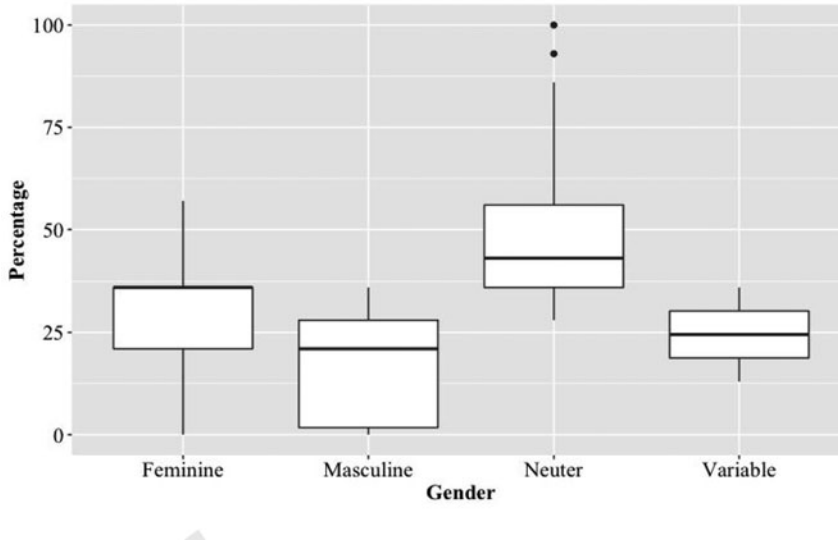

Figure 8. Adults: Gender Assignment in the Unproductive Condition

A paired t-test showed no significant difference between mean neuter assignment of monosyllabic and disyllabic nouns: $t(17) = -0.24$, $p = 0.81$. Figure 9 shows the distribution of gender assignment by syllable number.

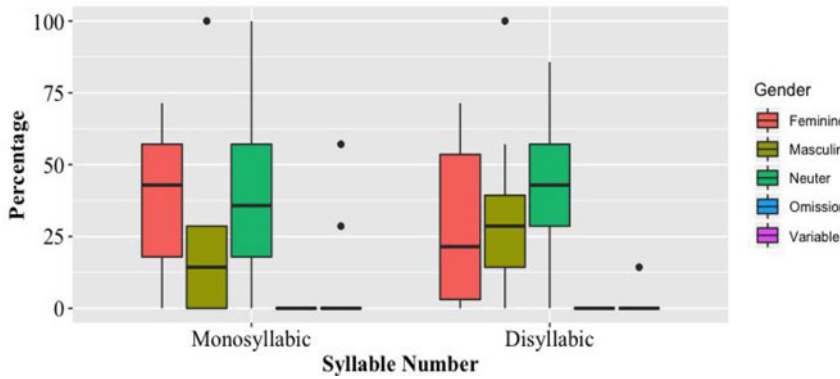

Figure 9. Adults: Gender Assignment and Syllable Number in the Unproductive Condition

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

Discussion

Overall, there were minimal differences between children's and adults' behavior in the task. However, adults assigned neuter significantly more frequently than children, as measured by a Welch's t-test: $t(31.54) = 2.39$, $p = .023$. There was no effect of age on children's performance. This suggests that a categorical distinction between productive and unproductive suffixes in Icelandic gender assignment can be made before the age of three on the basis of lexical experience, as predicted by the Tolerance Principle.

An alternative view of productivity

Productivity: categorical or gradient?

The Tolerance Principle predicted a categorical division between productive and unproductive processes in Icelandic gender assignment. However, a body of research has argued for an alternative view of productivity. On this view, productivity should be viewed and measured as a gradient phenomenon (Hay & Baayen, 2005; McClelland & Bybee, 2007). As a consequence, the difference between productive and unproductive patterns is not a categorical one and a pattern may be semi-productive.

A series of metrics to quantify morphological productivity at a scalar level have been proposed by Baayen and colleagues (Baayen, 1989; 1992; 1993). All of the metrics are centered around hapax legomena: namely, singleton words that appear precisely once in any given corpus. The general idea is that low token frequency should be a strong indication of productivity, given that lexicalized types in general have a higher token frequency than unlexicalized types.

The most studied metric proposed by Baayen and colleagues is P, which measures whether a given process is productive or not on the basis of token frequency. P is stated in (9), where n_1 represents the number of singleton words that a process applies to and N is the sum of the token frequencies of these items.

(9) $N = n_1/N$

The primary goal of P is to give a statistical measure of the probability of encountering new types (Baayen, 1993, p. 183). The larger the number of possible types, the more likely it is that they will not all occur in a given corpus or that some of them will occur only once.

A second metric, P*, compares one process against all other processes (Baayen, 1993). P* is stated in (10), where N_1 represents the total number of all singleton words that a process applies to.

(10) $P^* = n_1/N_1$

The primary goal of P* is to give a numerical estimate of the relative rate at which a category is expanding.

Baayen (1993, p. 194) proposed that P and P* should be viewed as two complementary measures; the primary use of P being to distinguish between productive and unproductive processes as such, while P* ranks proceses by degrees of productivity.[5]

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

Baayen's P and P* metrics were not explicitly designed to account for learning. Nevertheless, they have clear implications for learning. A comparison of the predictions of the Tolerance Principle and Baayen's metrics contributes to the dispute whether morphological learning involves detecting categorical or gradient patterns. Therefore, the three data sets presented in this paper were subjected to quantitative analyses using Baayen's P and P* metrics and their predictions evaluated against the empirical results.

Analysis using Baayen's P and P metrics*

Both P and P* are gradient measures of productivity, whereas the results of the elicited production task suggest that both children and adults make a categorical distinction between productive and unproductive suffixes in Icelandic gender assignment. This does not necessarily invalidate P and P* as quantitative measures. For instance, it is conceivable that there exists some quantitative threshold value that can be used to define productivity or absence thereof. How to construct such a threshold is beyond the scope of this paper. However, in the analysis below, I demonstrate important inconsistencies of the two metrics and discuss what gives rise to them.

Table 11 provides the results of a quantitative analysis using Baayen's P and P* metrics on Icelandic child-directed speech (adult), child naturalistic speech (child) and the SUBTLEX corpus. The denominator of P was the total number of tokens that take a particular suffix. The denominator of P* was the sum of all singletons attested for each gender.

There were two major types of inconsistencies in the values of the measures. First, P yielded radically different values depending on the corpus size due to its reliance on token counts (see Bauer, 2001, p. 153 for similar concerns). As a result, productive suffixes could be assessed as less productive than unproductive suffixes. Bold font in Table 11 indicates values that predict the productivity of unproductive patterns.

P* ranked suffixes more accurately; i.e. *–r* and *–i* were predicted to be most productive of masculine and *–a* was predicted to correlate with high or semi-productivity of feminine. Still, the ranking of the productive suffixes was variable between the two corpora (e.g., the productivity of *–r* and *–i* to masculine). This is because the value of P* is dependent on type counts which may vary between suffixes from corpus to corpus. As a result, the prediction for gender acquisition is that children should treat these suffixes differently depending on their type counts. However, neither children nor adults made such a distinction between the three productive suffixes in the elicited production task. Instead, they made a categorical distinction between productive and unproductive suffixes which is unaccounted for on a gradient approach to productivity.

General discussion and conclusion

In this paper, I have presented an approach whereby gender acquisition is driven by a search for productive patterns. Prior accounts have proposed that transparency is predictive of children's behavior in gender acquisition. I argue that transparency is a direct reflection of productivity. As a consequence, I propose that the term transparency be replaced with productivity.

Typological research on gender systems has revealed a wide range of possible gender assignment patterns (Corbett, 1991; 2013). Therefore, a theory of gender acquisition is

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

Table 11. Quantitative Analysis of Adult, Child and SUBTLEX Corpora

	M				F				N			
Corpus	*–r*	*–i*	*–a*	*–ø*	*–r*	*–i*	*–a*	*–ø*	*–r*	*–i*	*–a*	*–ø*
Adult												
P	0.14	0.19	0.03	0.15	**0.21**	0.12	0.11	0.08	**0.88**	0.08	**0.9**	0.12
P*	0.38	0.49	0.01	0.13	0.1	0.03	0.53	0.34	0	0.03	0.53	0.34
Child												
P	0.02	0.05	**0.9**	0.54	**0.9**	**0.89**	0.2	0.35	**0.78**	0.25	0.27	0.35
P*	0.48	0.25	0.03	0.19	0.01	0.13	0.74	0.7	0.09	0.2	0.33	0.58
SUBTLEX												
P	0.04	0.05	0	0.02	0.03	**0.1**	0.05	0.03	0.28	0.22	0	0.04
P*	0.59	0.37	0	0.11	0.022	0.067	0.48	0.61	0.01	0.16	0	0.013

needed that can account for how children can detect any kind of gender assignment pattern; be it semantic, morphological or phonological.

The present theory offers a general approach to how children detect gender assignment patterns. I have shown how predictions can be made using corpora as an estimate of the child's lexical experience in gender acquisition. As a result, any generalization about gender assignment can be subjected to the kind of quantitative analysis, proposed here, to make testable predictions.

Prior accounts of learning have argued that children categorically follow patterns that are frequent in the input in either experimental or naturalistic settings (Hudson Kam & Newport, 2005; 2009; Newport, 2019). However, a learning account must also be able to explain why children fail to generalize categorically on the basis of high frequency forms. Roughly a third of all noun tokens in Icelandic are neuter. Neuter nouns are also statistically dominant in the class of nouns that lack an overt nominative singular suffix. Still, neuter was not consistently chosen in the Unproductive condition. The unproductivity of neuter was predicted by the Tolerance Principle due to the number of masculine and feminine nouns of the same pattern.

Results from artificial language learning studies have shown that children tend to regularize linguistic patterns in the input data, even when these patterns show variability or inconsistencies (Hudson Kam & Newport, 2005; 2009). Thus, children do not merely reproduce the input statistics. However, the same studies found a different behavioral pattern for adults. Unlike children, adults matched the token frequencies of linguistic patterns instead of producing them in a categorical fashion.

Children and adults's response patterns in the present study were strikingly similar. The main difference involved the choice of neuter in the Unproductive condition, where adults used neuter significantly more often than children. This may suggest that some adult participants were trying to match the input statistics. Prior studies have shown that adults use irregular forms more often than children in experimental settings (see e.g., Berko, 1958). The source of child and adult differences in experimental settings remains unclear. In the present study, however, differences were only apparent in the Unproductive condition.

The results of the present study suggest that learning involves forming type-driven generalizations. Many contrasting theoretical approaches have recognized the role of type frequency in productivity. However, the main point of contention has been the division line between productive and unproductive processes. For instance, Bybee's (1985) Network model argues against a categorical division between productive and unproductive processes. Instead, the degrees of productivity of both productive and unproductive processes are determined by their token frequencies. As we have seen, such an approach makes inaccurate predictions with respect to Icelandic gender assignment. Baayen's approach is in the same gradient spirit and both types and tokens are made use of in his productivity calculations.

The empirical results presented in this paper do not support a gradient view of productivity: There were no differences in the degrees of productivity of the three suffixes in the Productive condition. In spite of statistical dominance, neuter was not consistently chosen in the Unproductive condition. Rather, the absence of a default gender manifested itself in inter-and intra-speaker variation. Hence, productivity resulted in categorical, uniform responses, whereas absence thereof resulted in inconsistency and differences in response patterns.

https://doi.org/10.1017/S0305000920000732 Published online by Cambridge University Press

Notes

[1] The term RULE is used in an atheoretical way in this paper and is compatible with other related terms such as PATTERN, REGULARITY or SCHEMA. On the present approach, rule formation is a consequence of a search for productive patterns in language acquisition. The author makes no commitment as to how rules discussed in this paper should be formulated or represented in theoretical terms.

[2] There exist two other correlations between nominative singular forms and gender assignment in Icelandic. Namely, nouns that end in either *–ing* or *–un* are invariantly feminine. However, only five noun types with *–ing* and two with *–un* were encountered in a corpus of child-directed speech (Sigurjónsdóttir, 2007). It is, therefore, a possibility that these patterns are not frequent enough to be detected by young children in gender acquisition.

[3] The majority of nouns in this class have an /u/ inserted between the suffix *–r* and *–i*. This is standardly assumed to be the result of an epenthesis rule (Thráinsson, 2017). In other words, the epenthesis is a purely phonological process, independent of gender assignment: that is, triggered automatically under suffixation.

[4] In linguistic research, default forms are expected when agreement is inert like, for instance, in the case of clausal subjects. However, it is at present unclear what role such forms play in the acquisition of gender assignment rules. For instance, Tsimpli and Hulk (2013) pointed out that children acquiring Dutch and Russian, over-generalize masculine despite that theoretically neuter has been claimed to be the default in both languages.

[5] Baayen has proposed additional metrics to address some concerns raised by his critics, but discussing them specifically is beyond the scope of this paper. The later metrics introduced by Baayen all rest on the same theoretical assumptions.

References

Albright, A. (2003). A quantitative study of Spanish paradigm gaps. In G. Garding & M. Tsujimura (Eds.), *WCCFL Proceedings, 22*, 1–14.

Aronoff, M. (1976). *Word formation in generative grammar*. MIT Press.

Aronoff, M. (1994). *Morphology by itself: stems and inflectional classes*. MIT Press.

Baayen, H. (1989). *A corpus-based approach to morphological productivity: statistical analysis and psycholinguistic interpretation* (Doctoral dissertation).

Baayen, H. (1992). Quantitative aspects of morphological productivity. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 1991* (pp. 109–149). Dordrecht: Springer.

Baayen, H. (1993). On frequency, transparency and productivity. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 1992* (pp. 181–208). Dordrecht: Springer.

Baronian, L., & Kulinich, E. (2012). Paradigm gaps in whole word morphology. *Irregularity in morphology (and beyond). Studia typologica*, *11*, 81–100.

Bauer, L. (2001). *Morphological productivity*. Cambridge University Press.

Berko, J. (1958). The child's learning of English morphology. *Word*, *14*, (2–3), 150–177.

Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. John Benjamins.

Clahsen, H. (1999). Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, *22*, 991–1069.

Clark, E. (1985). The acquisition of Romance, with special reference to French. In D. I. Slobin (Ed.), *The cross- linguistic study of language acquisition. Vol. 1. The data*. Hillsdale, NJ: Erlbaum.

Corbett, G. (1991). *Gender*. Cambridge University Press.

Corbett, G. G. (2013). Systems of gender assignment. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology.

Dabrowska, E. (2001). Learning a morphological system without a default: The Polish genitive. *Journal of Child Language*, *28*(3), 545–574.

Dabrowska, E. (2005). Productivity and beyond: mastering the Polish genitive inflection. *Journal of Child Language*, *32*(1), 191–205.

Fanselow, G., & Féry, C. (2002). Ineffability in grammar. In G. Fanselow & C. Féry (Eds.), *Resolving conflicts in grammars: optimality theory in syntax, morphology and phonology* (pp. 265–307). Hamburg: Helmut Buske Verlag.

Gorman, K., & Yang, C. (2018). When nobody wins. In F. Rainer, F. Gardani, H. C. Luschützky & W. U. Dressler (Eds.), *Competition in inflection and word formation* (pp. 169–193). Dordrecht: Springer.

Hall, K., Blake, A., Fry, M., Mackie, S., & McAuliffe, M. (2016). Phonological corpus tools, version 1.2. [computer program].

Halle, M. (1973). Prolegomena to a theory of word formation. *Linguistic Inquiry*, *4*(1), 3–16.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.

Hart, B., & Risley, T. R. (2003). The early catastrophe: The 30-million-word gap by age 3. *American Educator*, *27*(1), 4–9.

Hay, J., & Baayen, R. H. (2005). Shifting paradigms: gradient structure in morphology. *Trends in Cognitive Sciences*, *9*(7), 342–348.

Helgadóttir, S., *et al.* (2010). The Tagged Icelandic Corpus (MÍM). *Proceedings of the workshop on language technology of under-resourced languages* (pp. 67–72). Available at malheildir.arnastofnun.is/mim.

Henzl, V. M. (1975). Acquisition of grammatical gender in Czech. *Papers and Reports on Child Language Development*, 188–200.

Hernández-Pina, F. (1984). *Teorias psicosociolinguisticas y su aplicacion a la adquisicion del espanol como lengua materna*. Siglo XXI.

Hockett, C. F. (1958). *A course in modern linguistics*. Macmillan.

Horst, J. S., & Hout, M. C. (2016). The novel object and unusual (NOUN) database: A collection of novel images for use in experimental research. *Behavior Research Methods*, *48*, 1393– 1409.

Hudson Kam, C. L., & Newport, E. (2005). Regularizing unpredictable variation: the roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*(2), 151–195.

Hudson Kam, C. L., & Newport, E. (2009). Getting it right by getting it wrong: when learners change languages. *Cognitive psychology*, *59*(1), 30–66.

Iversen, R. (1922). *Norrøn grammatikk*. 7th edition. Tano.

Karmiloff-Smith, A. (1979). *A functional approach to child language. A study of determiners and reference*. Cambridge University Press.

Kiparsky, P. (1973). Elsewhere in phonology. In S. Anderson R., & P. Kiparsky (Eds.), *A festschrift for Morris Halle* (pp. 93–106). New York: Holt, Rinehart and Winston.

Kodner, J. (2019). Estimating child linguistic experience from historical corpora. *Glossa*, *4*(1), 122, 1–14.

Kvaran, G. (2005). *Orð. Handbók um beygingar-og orðmyndunarfræði* [*Word. A Handbook of Icelandic inflection and derivation*.] Almenna bókafélagið.

Levy, Y. (1983). The acquisition of Hebrew plurals: the case of missing gender category. *Journal of Child Language*, *10*, 188–200.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, *18*(3), 193–198.

Linaza, J., Sebastián, M. E., & del Barrio, C. (1981). Lenguaje, comunicación y comprensión. La adquisición del lenguaje. *Monografía de Infancia y Aprendizaje*, 195–198.

Mariscal, S. (2008). Early acquisition of gender agreement in the Spanish noun phrase: starting small. *Journal of Child Language 35*, 1–29.

McClelland, J. L., & Bybee, J. (2007). Gradience of gradience: a reply to Jackendoff. *The Linguistic Review*, *24*(4), 437–455.

Mills, A. (1986). *The acquisition of gender: a study of English and German grammatical development*. Berlin: Springer.

Newport, E. (2019). Children and adults as language learners: rules, variation and maturational change. In B. Landau (Ed.), *Topics in cognitive science*, 1–17.

Noreen, A. (1903). *Altisländische und Altnorwegische Grammatik*. Max Niemeyer.

Orgun, C. O., & Sprouse, R. (1999). From Mparse to control: deriving ungrammaticality. *Phonology*, *20*, 191–224.

Pertsova, K. (2005). How lexical conservatism can lead to paradigm gaps. In J. Heinz., A. Martin and K. Pertsova (Eds.), *UCLA Working Papers in Linguistics 11: Papers in Phonology 6*. UCLA Linguistics Department, Los Angeles, 13–30.

Pérez-Pereira, M. (1991). The acquisition of gender: what Spanish children tell us. *Journal of Child Language*, *18*, (3), 571–590.

Pind, J. (1991). *Íslensk orðtíðnibók [Frequency in Icelandic]*. Orðabók Háskólans.

Pinker, S., & Prince, A. (1994). Regular and irregular morphology and the psychological status of rules of grammar. In S. D Lima., R. L. Corrigan & G. K., Iverson (Eds.), *The reality of linguistic rules*, pages (pp. 230–251). Amsterdam: John Benjamins.

Plunkett, K., & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perception: implications for child language acquisition. *Cognition*, *38*(1), 1077–1106.

Rodina, Y., & Westergaard, M. (2012). A cue-based approach to the acquisition of grammatical gender in Russian. *Journal of Child Language*, *39*, 1077–1106.

Rodina, Y., & Westergaard, M. (2013). The acquisition of gender and declension class in a non-transparent system: monolinguals and bilinguals. *Studia Linguistica*, *67*, 47–67.

Rodina, Y., & Westergaard, M. (2015). Grammatical gender in Norwegian: language acquisition and language change. *Journal of Germanic Linguistics*, *27*(2), 145– 187.

Schuler, K., Yang, C., & Newport, E. (2016). Testing the Tolerance Principle: children form productive rules when it is more computationally efficient to do so. In *The 38th Cognitive Society Annual Meeting*, Philadelphia, PA.

Sigurjónsdóttir, S. (1991). *Interrogative Sentences in the Language of Two Icelandic Children* (MA thesis).

Sigurjónsdóttir, S. (2007). *The Fia corpus*. University of Iceland, Reykjavík.

Slobin, D. (1977). Language change in childhood and history. In J. Macnamara (Ed.), *Language Learning and Thought*. New York: Academic Press.

Steinmetz, D. (1985). Gender in German and Icelandic: inanimate nouns. In J. Faarlund (Ed.), *Germanic Linguistics. Papers from a symposium at the University of Chicago*. Bloomington: Indiana University Linguistics Club.

Szagun, G., Steinbrink, C., Franik, M., & Stumper, B. (2006). Development of vocabulary and grammar in young German-speaking children assessed with a German language development inventory. *First Language*, *26*(3), 259–280.

Thráinsson, H. (2017). U-umlaut in Icelandic and Faroese: Survival and death. In C. Bowern & R. Zanuttini (Eds.), *On looking into words (and beyond)* (pp.99–113). Berlin: Language Science Press.

Tomasello, M. (1992). *First verbs: a case study of early grammatical development*. Harvard University Press.

Tomasello, M. (2003). *Constructing a language*. Harvard University Press.

Tsimpli, I. M., & Hulk, A. (2013). Grammatical gender and the notion of default: insights from language acquisition. *Lingua 137*, 128–144.

Unsworth, S., & Hulk, A. (2010). L1 acquisition of neuter gender in Dutch: production and judgment. *Proceedings of Generative Approaches to Language Acquisition 2009* (pp. 50–51).

Yang, C. (2005). On productivity. *Linguistic Variation Yearbook*, *5*(1), 333–370.

Yang, C. (2016). *The price of linguistic productivity*: *How children learn to break rules of language*. MIT Press.

Cite this article: Björnsdóttir SM (2021). Productivity and the acquisition of gender. *Journal of Child Language* **48**, 1209–1234. https://doi.org/10.1017/S0305000920000732