

Productivity of a Polish child's inflectional noun morphology: a naturalistic study

Grzegorz Krajewski · Elena V. M. Lieven · Anna L. Theakston

Received: 12 December 2010 / Accepted: 12 December 2010 / Published online: 24 February 2012
© Springer Science+Business Media B.V. 2012

Abstract In the current paper, we examine the degree to which a 2-year-old Polish-speaking child exhibits productivity in her use of noun morphology. Using densely collected naturalistic data (five recording sessions per week) we assess the range of noun inflections she produces, the degree of productivity in her use of individual nouns, and the contextual productivity in her use of individual inflections. We adopt careful controls to allow comparison between the child's noun use and that of her caregiver. Our data show that although the child uses the same range of noun inflections as her mother, she shows a much more limited productivity in her use of both individual nouns and individual inflections with respect to their contexts of use. We discuss the results in the light of two different theoretical approaches to inflectional morphology: the usage-based, schema approach and the rule-based approach.

The notion that morphological processes can be explained by quasi-algebraic rules operating on symbolic units stems from the generativist approach to language and the nativist view of its acquisition on the one hand (Chomsky 1995), and from a particular approach to human cognition in general, the Computational Theory of Mind, on the other (Pinker 1997), according to which human cognition is essentially a system of rules, working in such an algorithmic way.¹ From this perspective,

¹ Although formal approaches to language can be traced back even to the nineteenth century, Chomsky's proposal not only exceeded them in formality, but also made a direct attempt to connect linguistics with psychology.

G. Krajewski (✉) · E. V. M. Lieven · A. L. Theakston
School of Psychological Sciences, University of Manchester, Manchester M13 9PL, UK
e-mail: krajewski@manchester.ac.uk

E. V. M. Lieven
Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

productive use of a noun inflection requires (1) the knowledge of an appropriate rule determining which ending should be attached to a particular noun stem in a particular grammatical case (strictly speaking, case-number combination; see below), (2) the ability to identify the stem as appropriate for a given rule (e.g., based on grammatical gender or phonological features), and (3) the ability to identify a given syntactic context as requiring a particular case. Since morphological rules (just as any other grammatical rules) operate on symbols, both 2 and 3 entail symbolic knowledge of grammatical categories like case or gender. Once a child has all the necessary knowledge, a given inflection is fully productive. However, since even simple morphological systems have parts that preclude rule-like descriptions (e.g., the past tense for irregular verbs in English), proponents of the rule-based approach allow an alternative route, with irregular forms rote-learned and stored in memory.

The Network Model (Bybee 1985, 1995; also Köpcke 1998) applies principles of the usage-based view of language use and acquisition to morphology (Tomasello 2003). According to this model, the development of morphology does not lead to the formulation of abstract symbolic rules but rather results in the gradual emergence of lexically based schemas of varying scope. When a child already stores a number of word forms sharing paired form and meaning similarities in her lexicon (e.g., a number of nouns sharing a particular ending marking the same inflection), she will begin to develop a schema, in which what was common to the original items remains specified, whereas in the place of varying material (roughly speaking, stems) an abstract slot emerges. Such a schema then works as a productive morphological pattern, into which a new lexical item can be accommodated (e.g., a noun can be inflected, even if the child has not heard a given inflectional form of that noun before). A central feature of this model, differentiating it from a rule-based approach, is that it presupposes partial and gradually increasing productivity. How productive a schema is, i.e., how 'inclusive' for novel stems its slot is, depends on the variability of the items already accommodated. The more items a given schema is based on (the greater its type frequency), the more easily it will be extended to new items. However, the more similar the existing exemplars are to each other, the more difficult it will be to extend the schema to different items. At the same time, newly accommodated items add on to the schema specification: entrenching the common parts and reinforcing the underspecification of the differing ones. Therefore only over the course of development will the inflectional schemas reach an adult-like state (for other accounts of language development seen as a phenomenon extended in time, see Berman 2004; Ravid 2004).

There is a large body of research on the development of the English past tense, with some researchers claiming that productive use of the regular *-ed* ending results from the application of a symbolic rule, with irregular inflections stored in memory as lexical entries (e.g., Pinker and Ullman 2002), while others argue for a single schema-based mechanism explaining both parts of the system (e.g., Marchman 1997; Maslen et al. 2004; Matthews and Theakston 2006). However, focusing on English verb inflections, when investigating the development of inflectional morphology in children, seems unfortunate, since the system, with only three regular endings (*-s*, *-ed*, *-ing*) and a limited number of inflectional forms per verb, lacks

complex patterns of allomorphy and syncretism, which are intrinsic to many inflectional systems. Moreover, the sharp distinction between regular and irregular past tense forms is paralleled by equally sharp differences in structure and frequency: the regular part of the system consists of a majority of verb types, and they all—irrespective of their form—attach *-ed*; the irregular part comprises a small set of highly frequent verb types, clustered into formally similar subsets, within which past tense forms result from particular alternations of the basic form. Thus, the very same regular/irregular distinction, which proponents of the rules approach derive from the distinction between grammar and lexicon in the human mind, from the schema-based perspective appears as an emergent result of frequency and similarity factors.

A complex system like Polish noun inflections seems more appropriate to study morphological development. Nouns in Polish attach endings that mark a combination of number (singular or plural) and case. Traditionally, there are seven case categories distinguished (nominative, accusative, genitive, dative, locative, instrumental, and vocative); Table 1 lists the main contexts of use for each of them: precise definitions of the case categories are difficult, if not impossible, due to fuzzy differences between the contexts across them combined with a substantial diversity of such contexts within categories (Dąbrowska 1997). The same case-number combination may be marked by different endings with a complex set of criteria governing the proper choice (grammatical gender, semantics, phonology) and in some cases there are no clear criteria at all; the same ending may mark different inflections (i.e., case-number combinations), sometimes for different classes of nouns, sometimes within the same class of nouns. Thus, a single noun may have up to 11 different inflectional forms, and the fact that attaching some endings to some stems triggers various stem alternations only adds to the complexity (for an overview of the morphology of Polish nouns, see Dressler et al. 1997).

Table 1 Polish cases, their main uses and typical endings (main uses adapted from Dąbrowska & Szczerbiński, 2006)

Case	Main uses	Singular endings	Plural endings
Nominative	Citation form; subject; subject predicative; form of address	\emptyset , <i>-o</i> , <i>-e</i> , <i>-ę</i> , <i>-a</i>	<i>-owie</i> , <i>-y/-i</i> , <i>-y/-i</i> , <i>-e</i> , <i>-a</i>
Accusative	Direct object; in certain PPs	\emptyset , <i>-a</i> , <i>-o</i> , <i>-e</i> , <i>-ę</i>	<i>-ów</i> , <i>-y/-i</i> , <i>-e</i> , <i>-a</i>
Genitive	Direct object of a negated verb; modifier of an NP; object of certain verbs	<i>-a</i> , <i>-u</i> , <i>-y/-i</i>	\emptyset , <i>-ów</i> , <i>-y/-i</i>
Dative	Indirect object; experiencer	<i>-owi</i> , <i>-u</i> , <i>-e</i> , <i>-y/-i</i>	<i>-om</i>
Locative	In certain PPs	<i>-e</i> , <i>-u</i> , <i>-y/-i</i>	<i>-ach</i>
Instrumental	Subject predicative; object of certain verbs	<i>-em</i> , <i>-ą</i>	<i>-ami</i>
Vocative	Form of address	<i>-e</i> , <i>-u</i> , <i>-o</i> , <i>-e</i> , <i>-ę</i> , <i>-y/-i</i>	=NOM

Smoczyńska (1985), based on some older observational studies and her own corpus analyses, describes the typical development of the system. According to her, the accusative and genitive singular emerge as the earliest contrasts to the nominative singular, usually in the third month of the two-word stage, i.e., around the age of 1;6. These are quickly followed by the nominative and accusative plural, with the rest of the singular inflections emerging by the age of 2 and the remaining parts of the system in the third year of life. The emergence of this complex system seems not only early and fast, but also relatively error-free: most endings are used correctly from the onset of a given category and several endings marking the same case appear simultaneously. Dąbrowska (2001), applying some productivity² measures (provision rates, overgeneralisation rates) to the same data and focusing on genitive endings, confirms the early acquisition of the endings and a very low rate of overgeneralisation errors for all of them (both Smoczyńska and Dąbrowska report only one case of persistent overgeneralisation: substituting *-ów* for \emptyset in the genitive plural).

This course of development might be accounted for by a rapid formation of symbolic rules. However, the fact that it is not always possible to provide precise criteria for the proper choice of an ending (e.g., for masculine inanimate nouns in the genitive singular) calls for some kind of a dual-mechanism account, with those “irregular” inflections stored in the lexicon. There are three reasons though to question such a possibility. First, it is difficult to perceive how a dual-mechanism model would account for switching between rule-governed and “irregular” (stored in memory) inflections and current empirical findings speak against it (Krajewski, Theakston, Lieven, & Tomasello, in press). Second, Dąbrowska (2001, 2004) reports the same gradual course of development for supposedly “irregular” and “regular” endings and no evidence for greater productivity of the latter. Third, symbolic rules require pre-given and clear-cut case categories to operate on. The prototype structure of use of Polish cases and fuzzy boundaries between them constitute thus another challenge for a rule-based approach.

The early and simultaneous use of noun inflections might be better accounted for by emerging schemas. In this view, it would be a consequence of gradual lexical learning. First occurrences of case-number combinations would simply be instances of producing rote-learned unanalysed inflectional noun forms and a child would only gradually learn to extend any given ending to a growing number of other nouns, thus accounting for the slow overgeneralisation rate reported in the literature. This process of abstraction would be guided by type and token frequency as well as similarity of form: the role of these factors has been shown in various experimental studies (Dąbrowska 2005, 2006; Dąbrowska and Szczerbiński 2006; Krajewski et al. in press; for a usage-based account of the development of Polish noun inflections based on naturalistic data, see Dressler et al. 1996).

Such a lexically restricted use of inflections in young children has been found in naturalistic studies on Italian (Pizzuto and Caselli 1992), Brazilian Portuguese

² In this paper we use the term *productivity* in its developmental psychological sense, referring to a child's (or an adult's) ability to use inflections, which goes beyond mere retrieval of rote-learned forms (see also Stephany and Voieikowa 2009). When we talk about *emergence* of inflections etc., we mean their first occurrence.

(Rubino and Pine 1998), Spanish (Aguado-Orea 2004; Mueller Gathercole et al. 2002), Croatian (Kovačević et al. 2009), and Palestinian Arabic (Ravid and Farah 2009). However, since inflectional schemas are pairings of form and meaning (which makes them a special case of grammatical constructions within the usage-based approach to language; Goldberg 2005), their slow emergence should be evident in both those aspects. Thus, just as gradual abstraction of form specification of a schema predicts a slow expansion of its lexical scope, gradual abstraction of its meaning specification should predict a narrow at first and slowly expanding scope of its contexts of use. This could eventually lead to the emergence of abstract yet prototype-based case categories. Whether the early use of inflections is contextually restricted has not been, to the best of our knowledge, studied so far.

Previous naturalistic studies of Polish (Dąbrowska 2001, 2004; Dressler et al. 1996; Smoczyńska 1985) used diary data gathered from a number of children. While having more than one child to base one's analyses on is an obvious advantage, the samples for each child were not gathered in any systematic way, with clear differences in quality between the children, and they were often too sparse at a particular point of development to allow reliable conclusions (e.g., Dąbrowska 2004). Here, we use a corpus of a single child densely sampled (five recording sessions per week) from 2;0 to 2;1, a period critical for the development of noun inflections. This is the first such dense dataset available in Polish.

Study 1 offers a preliminary analysis of the noun inflections used by the child. Its main aim is to provide an initial description of the corpus and to set the scene for the proper analyses in Studies 2 and 3. The aim of Study 2 is to test whether the use of noun inflections in a child's spontaneous speech sample is restricted lexically, and the aim of Study 3 is to test whether it is restricted contextually. Importantly, both Studies 2 and 3 rely on statistically comparing the child sample to an adult speaker sample from the same corpus (the mother), thus deriving the standard for evaluating the child from her input language and avoiding arbitrary criteria. Since adult use of inflections may appear more productive simply because of larger vocabulary size, knowledge of more inflections, longer MLU, or broader range of available contexts, all those factors are carefully controlled. If nonetheless there is a difference between the use of noun inflections by the child and by the mother, in terms of the possible nouns they are used with and/or their potential contexts of use, it would provide some support for a schema-based model.

Study 1

Data and coding

The corpus was gathered from a two-year-old girl (Marysia), coming from a typical monolingual Polish family from the Gdańsk region.³ She was followed for six weeks. There were 30 audio recording sessions (five per week), each taking place at home and

³ The corpus was collected by Ewa Dąbrowska and funded by the Max Planck Institute for Evolutionary Anthropology, Leipzig.

lasting about an hour, with the girl interacting mostly with family members during everyday play activities (her mother did the recordings). The first session took place when she was 2;0.3, the last one when she was 2;1.12. Thus, the sampling covers the period when most noun inflections should have already appeared (Smoczyńska 1985). At the same time, approaching their development from a schema-based perspective, one would expect productive patterns only to begin to emerge. There were five sessions each week on average, which gives a sample constituting roughly 7% of all the verbal activity of the child in the recording period (assuming she is active 10 hours a day). Such dense sampling is still rare in the child language field (Tomasello and Stahl 2004) and there are only a few (but increasing number of) studies based on dense datasets of English speaking children (Cameron-Faulkner et al. 2007; Dąbrowska and Lieven 2005; Kirjavainen, Theakston, and Lieven 2009; Kirjavainen, Theakston, Lieven and Tomasello 2009; Lieven et al. 2003; Maslen et al. 2004; Theakston and Lieven 2008) and of a German speaking child (Abbot-Smith and Behrens 2006). Thus the analyses presented in this paper make an important addition to the field.

Table 2 presents preliminary corpus counts. All imitations, repetitions, and quotations were excluded both from the child's and the mother's speech. All utterances were morphologically coded (in CHAT format; MacWhinney 2000); in particular, for each noun form its lemma and case-number combination marked by its ending were coded. When resolving any coding issues, to minimise the arbitrariness of our decisions, we prioritised the criterion of form, above syntax, semantics, or discourse context. Thus, two homonymous nouns would be coded as one noun type, unless there were differences between their inflectional paradigms (i.e., not all their inflectional forms were homonymous; e.g., the same noun might refer to an object and be a proper name, with its accusative form depending on the use; hence the two uses would be coded as different types, e.g., "kapturek" [*hood-DIM*] versus "Kapturek" [as in *Little Red Riding Hood*], with the accusative "kapturek" versus "Kapturk-a" accordingly); parts (including nouns) constituting prefabricated chunks or idioms were coded separately (e.g., *for example* and *good night* were not coded as single items; instead, *example* and *night* were counted as occurrences of these particular noun types, in the accusative and nominative singular respectively). Similarly, the case-number combination of a noun form was coded based on its ending, rather than on the requirement of its syntactic context, unless attaching the ending resulted in a form which does not exist in adult Polish (including its non-standard varieties), in which case the form would be coded as "non-adult" (e.g., when the child produced "do fontann-ę" [*to fountain-ACC:FEM*],

Table 2 Basic corpus characteristics

Participant	Utterances	MLU	Utterances containing nouns	Noun tokens
Marysia	17,056	2.733	6,524 (38.25%)	7,224
Mother	15,157	4.925	6,885 (45.42%)	8,606
Total ^a	37,263	3.902	15,906 (42.68%)	18,991

^a Other participants also contribute to the data but their contributions are significantly smaller (e.g., Father's total number of utterances, which is third largest after the Mother's, amounts to 3,443).

using the accusative instead of the genitive required by the prepositional *to* phrase, it was coded as an accusative use of that noun, but when she produced “chlebk-ą” [*bread: DIM-INSTR: FEM*] with a feminine instrumental ending instead of a masculine genitive, it was coded as a morphological error). The context, syntactic or otherwise, had to be taken into account, when coding syncretic forms (i.e., resolving which case a given ending marks in a given context). There were 8346 (44%) such forms in the whole corpus and we attempted to disambiguate as many of them as possible by reference to the linguistic context, which potentially led us to credit the child with more productivity than she actually had. Nevertheless we had to exclude 344 (2%) ambiguous forms.

The disambiguation was conducted by the first author and, given its subjectivity, the reliability of the whole process was checked. Four hours (out of a total of 30) were selected at random and within this subset all syncretic noun forms were disambiguated by a second coder (a native speaker of Polish holding an MA in Polish language). Approximately 1/8 of this material was used for training purposes and the reliability check was conducted on the remaining 7/8. There were 851 syncretic noun forms (10.20% of all 8346) and the Cohen's Kappa coefficient was 0.88 indicating a very high rate of agreement.

Results

Table 3 presents frequency counts for noun forms in the two speakers' samples broken down according to the case-number combinations and endings. Two facts are worth noticing. First, all case-number combinations are attested in both speakers' samples⁴ and both speakers use roughly the same endings in all those combinations. There is only one ending in one case-number category present in the mother's sample that was unattested in the child's speech (the accusative singular *-ą*, used in “pani-ą” [*lady-ACC*]). However, the ending is very infrequent in the mother's speech as well (type/token = 1/3) and there is also one ending unattested in the mother's sample (the nominative plural *-owie*, which is rare in the child's speech too, type/token = 1/5, used in “pastuszk-owie” [shepherds]). Thus the data so far support the claim of the early emergence of noun inflections in Polish children. Second, although the numbers are higher in the adult sample, the distributions of noun inflections remain remarkably similar in both samples. Indeed, for type and token frequencies Pearson's $r = 0.99$ and $r = 0.97$ respectively, with the correlation of type/token ratios, $r = 0.39$, $p < 0.011$. Interestingly, these frequency counts resemble the typical order of emergence reported in the literature (the accusative, genitive, and locative singular being the earliest contrasts to the nominative, followed by the nominative and accusative plural, followed by the rest of the system; Dressler et al., 1996; Smoczyńska, 1985), although the dense corpus allowed us to sample in the child's speech, just after her second birthday, some occurrences of all

⁴ The total number of combinations considered in this paper is 13, rather than 14 (7 cases \times 2 numbers). The vocative plural is only a theoretical possibility and was not included in coding, since (a) it is syncretic with the nominative across all possible nouns, (b) in Polish the vocative in general is infrequent, with the nominative being often used instead even in singular contexts, and (c) it is immanently difficult for single-word utterances to tell the nominative and the vocative contexts apart.

Table 3 Frequency distribution of noun inflections in the child's and in the mother's sample (sorted by the type frequency of case-number combinations and of endings within them in the child's sample)

	Inflection	Marysia		Mother	
		Type	Token	Type	Token
NOM:SG	∅	258	1,260	341	1,315
	-a	246	1,186	306	1,818
	-o	42	136	55	187
	-e	8	14	26	63
	-i/-y	5	33	5	45
	-ę	4	10	4	10
	Total	563	2,639	737	3,438
ACC:SG	-ę	129	420	179	574
	∅	100	262	157	519
	-a	51	155	63	149
	-o	30	78	38	106
	-e	5	8	18	43
	-ą	0	0	1	3
	Total	315	923	456	1,394
GEN:SG	-a	100	243	153	306
	-i/-y	89	327	136	437
	-u	29	90	52	126
	Total	218	660	341	869
LOC:SG	-u	59	130	96	260
	-e	58	121	104	278
	-i/-y	14	31	21	56
	Total	131	282	221	594
NOM:PL	-i/-y	100	168	131	308
	-e	14	52	21	111
	-a	9	16	16	36
	-owie	1	5	0	0
	Total	124	241	168	455
ACC:PL	-i/-y	79	213	117	397
	-a	16	46	22	64
	-e	5	18	15	29
	Total	100	277	154	490
INSTR:SG	-em	42	70	72	145
	-ą	23	37	43	91
	Total	65	107	115	236
VOC:SG	-u	34	210	38	268
	-o	9	25	20	32
	-e	2	23	7	27
	Total	45	258	65	327

Table 3 continued

	Inflection	Marysia		Mother	
		Type	Token	Type	Token
GEN:PL	∅	22	30	40	84
	-ów	16	38	43	121
	-i/-y	5	11	17	39
	-u	1	1	1	11
	Total	44	80	101	255
DAT:SG	-owi	15	24	14	36
	-i/-y	5	10	9	72
	-e	2	2	11	24
	-u	2	5	2	3
	Total	24	41	36	135
LOC:PL	-ach	20	32	42	53
INSTR:PL	-ami	10	13	40	54
DAT:PL	-om	1	1	3	6

the infrequent plural inflections, which were said to typically emerge during the third year of life.

It might be concluded that on the surface level the child's use of noun inflections is very much adult-like. However, we identified a number of non-adult forms, i.e., overgeneralisations of some endings. There were 36 occurrences of them (distributed over 27 types), amounting to 0.64% of all noun tokens in the child's sample. This number seems negligibly small and it might be tempting to dismiss it as noise in the data, since there are six non-adult tokens (0.07% of the total, mostly direct imitations of the child) in the mother's speech as well. However, the difference between the two percentages is highly significant by a chi-squared test, $\chi^2(1) = 36.31$, $p < 0.001$, and all but one non-adult form in the mother's speech are direct imitations of the child. Hence, the child does sometimes overgeneralise noun inflections. The very low rate of such overgeneralisations is in accord with previous findings. However, two interesting facts should also be noted. First, most overgeneralisations found in the child's speech could happen in both directions, i.e., if one ending was overgeneralised for another one, the latter might have been overgeneralised for the former as well (e.g., "kot-owi" instead of "kot-u" [*cat-DAT*], but "Fred-u" instead of "Fred-owi" [*Fred-DAT*]). Second, for at least 11 errors intended adult forms were attested in the child's sample too (some errors excluded the possibility of identifying the intended form). There is thus no evidence of any consistent and persistent overgeneralisations.

Obviously, this short account of overgeneralisations does not amount to a comprehensive analysis of errors in the child's use of noun inflections. Particularly interesting might be to investigate her failure to use appropriate forms in given

contexts (irrespective of whether she used different valid case forms or non-adult forms instead). This would go well beyond the scope of this paper requiring a whole new study. However, though we cannot give specific numbers, our impression, based on the analyses of the corpus we have done so far, is that the overall rate of such errors would not be very large either. Moreover, there is some experimental evidence that 2-and-a-half-year olds do not make such errors even when inflecting novel nouns (rather, they would use a wrong gender ending or simply repeat the form they heard; Krajewski et al. in press).

To summarise this introductory description, the frequency distributions suggest adult-like use of noun inflections by the child, but the occurrence of non-adult forms distorts this picture. At the same time, there is no evidence for a rule-governed overgeneralisation of any inflectional pattern. Rather, the errors might have resulted from weak and not fully entrenched emerging schemas. Whatever the exact nature of the mechanism underlying the inflectional system is, the results thus suggest that its development has not yet reached the productive end-state. The following two studies are intended to methodically test this claim.

Study 2

The aim of this study was to evaluate how lexically diverse (or restricted) the child's use of noun inflections is. Measuring the lexical diversity of inflections (i.e., how freely they are used with different lexical types) in child language corpora, being intrinsically related to measuring morphological productivity, is particularly prominent within the usage-based framework, as this approach makes a crucial distinction between the first appearance of an inflection and its productive use. The latter develops only gradually as a function of (among other factors) the number of lexical types the inflection appears with. Therefore, the occurrence of a certain number of types in a given inflectional pattern (suggesting some lexical diversity of that pattern) has been used as an indirect measure of acquired productivity with the pattern (e.g., Dąbrowska 2004; Pizzuto and Caselli 1992). However, this measure is prone to a sampling bias: the more frequent an inflection is, the more likely it is to be sampled with a larger number of types. Also, the criterion, whatever the number of types one decides on, is completely arbitrary and does not guarantee that the inflection has been used productively, as it could be that all forms were merely retrieved from memory. To minimise the chance of this possibility, an additional criterion may be introduced: the lexical types used with a given inflection should also occur with a particular number of other inflections (so-called *contrastive use*; Mueller Gathercole et al. 2002). Pine et al. (1998) used a similar measure, calculating a ratio of the number of types used contrastively in a given inflection to the total number of types used in that inflection and testing the overlap statistically. Using a ratio and a statistical test avoids both an arbitrary cut-off point and a potential sampling bias, being fair to all inflections independent of their frequency.

Nevertheless, all the above-mentioned measures of productivity have in common the fact that they assess the lexical diversity of a morphological pattern by applying

some arbitrary criterion to it.⁵ Moreover, they presuppose an idealistic end-state: all nouns in adult speech, it is implicitly assumed, are equally likely to appear in all inflectional forms. Assessing productivity in child speech effectively means testing how far it is from that state. However, it is by no means clear that adult use of inflections represents this end-state. Use of certain lexical types may be limited to particular inflections and use of certain inflections may be limited to particular types. Also, even adult spontaneous speech is characterised by a substantial amount of repetitiveness and therefore in a limited sample adult speakers may not reveal their full productivity even with the least lexically restricted inflections. Productivity may not manifest itself in adult spontaneous speech in the way that has previously been assumed when assessing child language.

The way forward might be to compare the child's speech sample directly with the sample of a child directed speech provided by a caregiver. Aguado-Orea (2004), investigating the development of verbal morphology in Spanish, compared the average number of inflections per verb in a child's sample and in the parent's sample. Thus, the criterion for evaluating the child's use of inflections was derived directly from her input language, avoiding both the question of how much productivity to expect and the need to make any arbitrary decisions. This method cannot provide information on individual inflections, but as an overall measure of lexical diversity in their use—provided that other potentially biasing differences between a child and an adult speech are controlled—it might be considered a reasonable measure of a young child's productivity of an inflectional system.

Method

The method used in this study was based on Aguado-Orea (2004). We counted the number of inflections per noun type in the child's and the mother's sample and compared the two samples statistically, controlling three potentially distorting factors.

First, the difference in the vocabulary size might bias the comparison of the two speakers in a non-trivial way. On the one hand, the greater vocabulary size of the adult resulted in more opportunities to use the various inflections and consequently might inflate the average number of inflections per noun type in the adult sample. On the other hand, greater diversity of noun types used by the adult might result in fewer inflections per type. Thus, we included only the noun types attested in both speakers' samples. Having the same set of types in both samples allowed us to use a paired-samples statistical test, which in turn controlled lexical variation in the use of different inflections, i.e., the fact that different nouns are more likely to appear in different (number of) case forms.

Second, due to a greater mean length of utterance (MLU) the adult sample (number of noun tokens) was considerably larger than the child sample, which led to

⁵ Even Pine et al.'s ratio does not avoid arbitrariness when defining contrastive use. This becomes apparent when analysing morphological systems richer than the English past tense. When there is a greater number of available inflections, this raises the question of how many of them should be used with a given type to provide evidence of contrastive use (see Krajewski 2005).

the possibility of attesting fewer inflectional forms of the same nouns in the latter, simply due to the smaller number of noun tokens overall. Hence, for each noun type included we equalised the number of its tokens in both samples, randomly excluding tokens from the larger sample. To ensure stability of the results, the procedure of random reduction was reiterated 1,000 times.

Third, adult language might appear more productive because a mature speaker has more inflectional patterns at her disposal. Deeming the child as not using noun inflections productively only because she does not know as many of them as her mother would be unreasonable. Therefore, we included in the analysis only those inflections, which were attested in both speakers' samples.

A separate issue concerns what exactly should be considered as different inflections, which is important for both counting inflections and excluding the ones that appeared only in one speaker's sample. The answer is not straightforward, since the system is full of syncretisms and allomorphy, i.e., the same ending may mark several case-number combinations and several different endings may mark the same case-number combination. Taking allomorphy into account, we first conceptualised inflection as a particular ending marking a particular case-number combination. Thus, we counted an occurrence of a noun with a given ending only if the ending marking a given combination was attested in both samples. Conversely, a given ending was counted more than once if it was used to mark more than one case-number combination, provided that respective uses were attested in both samples. This way, it was possible that, for example, the child used a particular feminine noun in a genitive, dative and locative singular context, and was thus credited with three inflections for that noun, even though all three of them required the same syncretic form. This might result in an overgenerous assessment of the child's productivity, but it does not seem unreasonable since the child contrastively marked these inflections with other endings as well (e.g., using contrastively *-owi* in the dative singular and *-u* in the locative singular).

Nevertheless, we also ran a second, more conservative, comparison, taking into account the syncretism and counting only distinct forms of nouns instead of case-number combinations in which those nouns appeared. Still, the restriction on the forms to include held: only the forms resulting from the attachment of an ending attested in both speakers' samples counted. If the first, inflection-based, analysis indeed overestimates the child's productivity, we would expect an even greater difference between the child and the mother here, in the form-based analysis.

Results and discussion

There were 13,931 noun tokens extracted from the corpus, 8,321 from the mother's utterances and 5,610 from the child's utterances. The child's tokens were distributed among 907 noun types and the mother's tokens among 1,288 noun types. 734 of those types were shared by both speakers and thus included in the analyses. For each of the included types, the number of tokens in both speakers' samples was randomly equalised. Then both samples were compared on the number of inflections (inflection-based analysis) or distinct forms (form-based analysis) per noun type, with only inflections or, respectively, forms with endings occurring in both samples

included. To make sure that despite the random element in the procedure, the results were stable, the whole procedure was repeated 1,000 times. A set of Perl (version 5.8.8) scripts were used to extract and count data and an R (version 2.8.0) script was used to iterate the analyses, each time calculating statistics. Figure 1a and b presents relative frequency histograms of the number of inflections per noun type for the child and for the mother respectively, based on averaged counts obtained after applying the control procedures. Both distributions are highly skewed with around half of the noun types used with only one inflection and a further quarter with two inflections by both speakers. However, in both samples there are also noun types used with as many as eight and nine inflections. Figure 2a and b gives a similar picture for form-based counts.

The mean number of inflections per type was 1.86 in the child's sample and 2.00 in the mother's sample (calculated over all 1,000 comparisons; $SD = 0.007$ for the child, $SD = 0.011$ for the mother).⁶ Given the violation of the normality assumption, Wilcoxon signed-rank tests were run. The mean of the 1,000 z scores was 4.50 ($SD = 0.45$) and the mean of the exact p values was less than 0.001 ($SD < 0.001$). Thus for the inflection-based analysis, the differences between the two distributions were significant.

The mean number of distinct forms per type was 1.71 ($SD = 0.007$) in the child's sample and 1.79 ($SD = 0.01$) in the mother's sample. The mean of the z scores calculated for Wilcoxon signed-rank tests was 3.10 ($SD = 0.46$) and the mean of the exact p values was less than 0.006 ($SD = 0.01$). Thus, the differences between the distributions taken from the child's sample and those taken from the mother's sample, although smaller than in the inflection-based analysis, were still significant.

The results suggest that the mother was likely to use the same noun in a greater number of inflections (forms) than the child and that this could not be accounted for by the child having fewer inflections available, nor by her producing shorter utterances. This is a substantial difference between the two speakers' spontaneous use of noun inflections, indiscernible when merely looking at the provision of individual inflections. The fact that a number of confounding factors that could potentially contribute to such a difference were controlled strongly suggests that it was caused by an underlying difference in the speakers' knowledge of the inflectional system. That is to say, the child could not, at that stage of her language development, apply noun inflections as freely as an adult. Although this cannot be definitively determined from naturalistic data, she seemed to be beyond the stage of purely rote-learned forms (e.g., she used some non-adult forms). Nonetheless, having the same set of stems and the same repertoire of inflections as her mother, her inflectional ability made her combine at least some inflections with a limited number of stems as compared to her mother. This means that the underlying mechanism (at least for some inflections), although operating, was limited in scope.

Nonetheless, the results present an issue, potentially challenging for this conclusion, concerning the relatively small effect size. The two distributions did not

⁶ Standard deviations reported in parentheses refer to distributions of respective statistics across 1,000 re-samplings.

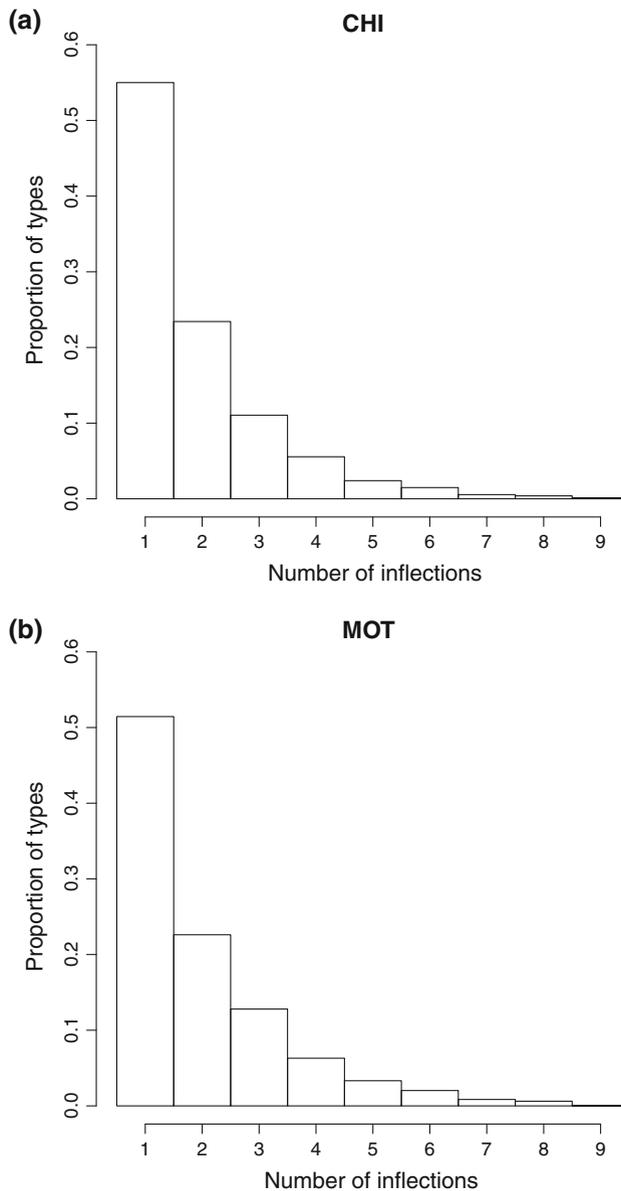


Fig. 1 Proportions of noun types occurring with different numbers of inflections **a** in the child's sample, and **b** in the mother's sample (averaged across 1,000 samples)

differ a lot from each other in both analyses (see Figs. 1, 3), the difference between means being on average 0.14 in the inflection-based analysis, and only 0.08 in the form-based analysis. There is no general agreement on measuring the effect size for non-normal data; nevertheless two different indices confirm the intuition, r (Field 2005) and A (Vargha and Delaney 2000): mean values calculated over 1,000

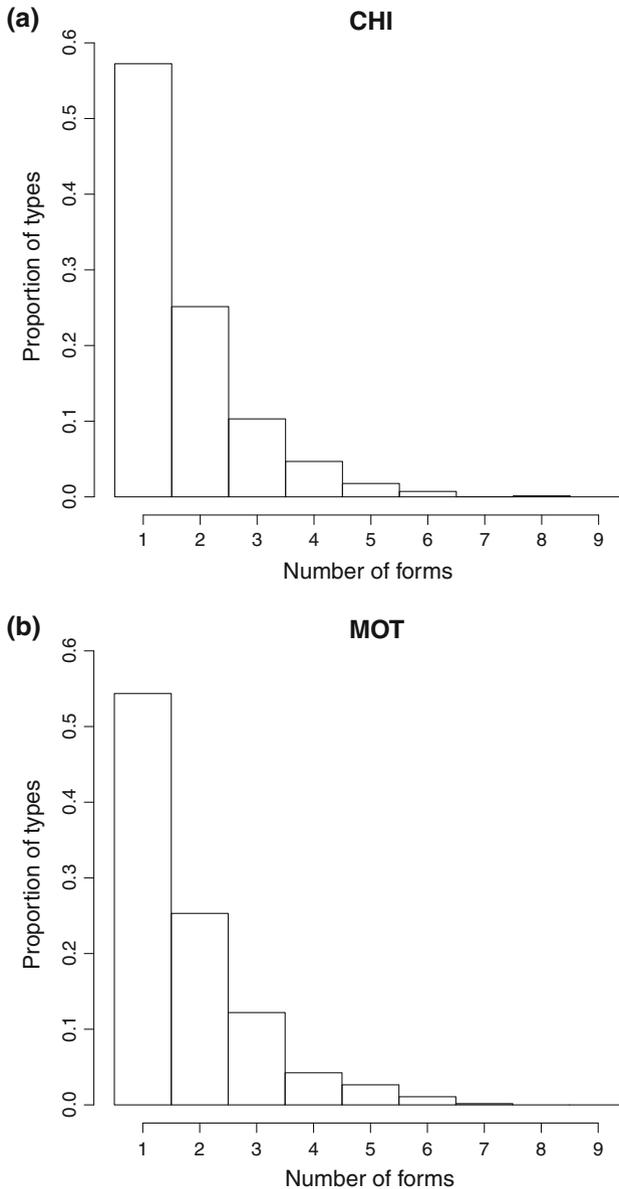


Fig. 2 Proportions of noun types occurring in different numbers of distinct forms **a** in the child's sample, and **b** in the mother's sample (averaged across 1,000 samples)

comparisons were $r = 0.12$, $A = 0.55$ in the inflection-based analysis, and $r = 0.08$, $A = 0.53$ in the form-based analysis.

Our contention is that this apparently weak effect size was imposed by the nature of the data, i.e., a relatively small range of possible values and highly skewed distribu-

tions, both stemming from the majority of noun types taking one or two inflections in the two speakers' samples and no noun types taking more than nine inflections. The skewness of distributions restricts the potential difference between central tendencies and the small range of values results in a difference score of 0 for many noun types, which is at least partly responsible for the small values of both effect size indices. Nevertheless, it should be noticed that there are no objective and generally applicable means of interpreting the magnitude of an effect size (various rules of thumb are often used) and each time the researchers have to decide for themselves which value they find satisfactory (Howell 2007). Given the constraining characteristics common to both distributions, it thus seems that one should have expected small values from the beginning. The important fact is that even with these restrictions the difference between the child and the mother (in terms of whole distributions, rather than just means) is still significant. Interestingly, contrary to the earlier concerns, the inflection-based analysis does not appear to overestimate the child's productivity. Counting syncretic forms as the same resulted in even smaller differences between the child and the mother, since it resulted in proportionally similar reduction of numbers in both samples. Importantly, it should be noticed that the high skewness of the mother's sample results from the fact that even she used the majority of nouns in a single inflection. This justifies our concern that the adult language is distant from the idealistic "full" productivity and supports the decision to compare the child with an adult.

To conclude, the study provides evidence that a two-year old child's use of Polish noun inflections is lexically restricted as compared to child directed speech. This finding is particularly interesting in the light of the fact that (1) the child was already using almost all inflectional endings available in the system and, at the same time, (2) their frequency distribution was similar to that found in an adult speaker without any ending being notably overgeneralised. It suggests that at this stage there were no highly productive mechanisms, e.g., rules, at work, but rather that the learning process was gradual and lexically based. However, focusing on the lexical aspect of using morphological patterns when investigating their productivity (i.e. how many lexical items they accommodate or how open they are to new ones) overlooks the equally important, particularly from a usage-based perspective, contextual aspect (i.e., how many different semantic or syntactic functions they serve or how easily they are extended to new ones). Study 3 addresses this issue.

Study 3

Although the contextual aspect of morphological development has not been studied extensively, Bybee's Network Model (Bybee, 1985, 1995) allows the formulation of certain predictions. In this approach, morphological productivity is accounted for by schemas, which gradually emerge based on commonalities of word forms stored in memory. Importantly however, the commonalities involve more than just shared features of form; they entail common *form-meaning pairings*. For example, storing a number of English word forms ending in *-s* may result in a plural schema being abstracted, but only if those word forms are all plural nouns. Similarly, a number of Polish noun forms sharing the same ending will give rise to a schema, if (and

because) the occurrence of that particular ending is related to (a) certain context(s) of use. In the end, morphological schemas are yet another kind of linguistic construction (Croft 2001; Goldberg 2005), i.e., pairings of form and meaning, where the form is partially specified (for Polish nouns, the ending attached to the noun stem, perhaps with a partial specification of the final consonant of the stem, as it sometimes conditions the choice of the ending) and the meaning identifies the context in which a given inflection is required.

Within the usage-based approach, schemas are assumed to reach adult-like productivity only gradually over the course of development. Early child schemas are based on a limited number of items, and thus show only limited productivity since the extent to which a schema is open to novel items depends on the variability of the items already accommodated. As the child learns more exemplars, these newly accommodated items add to the formal specification of the schema, entrenching the common parts and reinforcing the underspecification of the differing ones. It might be argued that a parallel process takes place with respect to developing the contextual specification of a schema. In order to acquire morphological productivity, children must (1) notice the endings as common parts of various noun forms used in the same contexts and, in parallel, (2) detect commonalities between different contexts requiring the same ending, in order to (3) develop productive schemas based around the endings in appropriate contexts. It may be argued thus that for a young child whose inflectional schemas have not yet achieved adult-like productivity, specifications of possible contexts of use, which are part of semantic specifications of the child's schemas, may be narrower than corresponding specifications in adult language. Hence the contexts of use of the endings in the child's speech may, as a result, be significantly restricted as compared to an adult's speech. The goal of the study presented here is to test this hypothesis.

We approached the problem of measuring contextual restrictiveness from a computational perspective. This avoids inevitable pitfalls and arbitrary decisions, when defining distinct contexts, judging similarity between them, choosing between semantic and syntactic features etc., and allows straightforward comparisons. Various studies have established that distributional properties of linguistic items, and patterns of their co-occurrence in particular, reflect similarities between them, which can be validated experimentally (Burgess and Lund 1997; Landauer and Dumais 1997). Those emerging similarities not only involve semantic features, but also more abstract, grammatical properties, both broad word class membership and fine-grained differences between inflectional forms (past-tense vs. past-participle in English), implying a blurred distinction between semantics and syntax. Hence, in this study, rather than defining contexts in some abstract syntactic way, we stayed on the lexical level and measured diversity of lexical items surrounding particular noun inflections. The assumption was that if, despite controlling the vocabulary and sample sizes, we found a difference it would imply an underlying difference in knowledge about potential contexts of use of the inflections (this mechanistic approach, however, is not free from its own pitfalls; see the discussion of results).

We employed a measure devised by McDonald and Shillcock (2001), stemming from Information Theory (Shannon 1948). The basic premise of the method is that the more restricted the contexts of use are for a given item, the more informative is

that item about its potential context. For example, the word *amok* is highly informative, since knowing about its occurrence we can be almost certain that the word *run* co-occurred with it. On the other hand, *run* can appear in a wide variety of linguistic contexts and consequently its occurrence does not convey much information about its potential context (almost any word may co-occur with *run*). The measure, which the authors call Contextual Distinctiveness (CD), quantifies this difference, with higher values indicating lower diversity of contexts. McDonald and Shillcock found a significant relationship between CD values and reaction times in a visual recognition task and argued that the amount of information brought by a word, which is what CD captures, affects the processing time.

Here, we calculated CD values for noun inflections, rather than individual noun types used with those inflections, and for each inflection we compared its CD value in the child's and the mother's speech samples. Assuming the hypothesis about the child's use of inflections being contextually restricted, we expected on average significantly larger CD values in her sample.

Method

To compare the diversity of contexts of use of noun inflections in the child's and the mother's sample, we pulled out all those speakers' utterances, which were at least two words long and contained a noun form other than a vocative or a non-adult form. We divided these utterances with respect to the inflection of the noun form (defined as a particular ending marking a particular case-number combination). For each inflection, we collapsed utterances across noun types, as we were not interested in the diversity of contexts in which any particular noun form might be used, but rather in the contexts of the inflection. However, we first excluded utterances with noun forms found in one speaker's sample only, in order to eliminate the effect of vocabulary size (a noun form not known to the child might contribute to a greater diversity of contexts of an inflection in the mother's sample). Also, the greater sample size, i.e., the overall larger number of noun tokens in the mother's sample, might result in a greater apparent diversity of contexts, thus we randomly excluded tokens to match their number in the child's sample. This random procedure was repeated 1000 times, to ensure stable and reliable results.

Noun inflections used by both speakers formed the set of *target items*, for which we wanted to calculate CD values in the two speech samples. First, we had to define the set of *context items*, which would be used in assessing the diversity of context, and we had to set the size of the *window of co-occurrence*, i.e., the maximum distance (counted in word forms) between a target and a context item for which the pair would be regarded as co-occurring. To define the set of context items, we included all word (surface) forms shared by both speakers, so that the set of context items used to calculate CD values was the same in both samples. This way we controlled a potential effect of vocabulary size. Focusing on surface forms means that different inflectional forms of the same word were treated as distinct items: they were counted independently; however, if a word was used by both speakers, but in different forms, it was not counted at all. We decided to treat different forms of a word separately, since the

grammatical form of a word might importantly add to the context, of which the word is a part. Moreover, in doing so, we avoided crediting the child with any abstract knowledge. Since the MLU in words was less than 2 in the child's sample, the window size was set to ± 1 word, i.e., only adjacent words were counted as co-occurring. Otherwise, the comparison with the mother might be biased.

CD is a relative entropy measure, a measure of dissimilarity of two probability distributions (MacKay 2003). For a set of context items (c), we will define their *prior* probability distribution, $P(c)$, as a distribution of their relative (i.e., divided by the total) frequencies in a corpus. The prior distribution tells us how likely it is to encounter these items in the corpus. The values of the prior distribution are obtained using the formula from Eq. 1, where $f(c_i)$ is the frequency of c_i in the corpus and the denominator is the total frequency of all n context items in the corpus.

$$P(c_i) = \frac{f(c_i)}{\sum_{j=1}^n f(c_j)} \quad (1)$$

A *posterior* probability distribution of the same set of context items, $P(c|w)$, will tell us how likely it is to encounter these items co-occurring (within a co-occurrence window) with a given target item (w). In other words, it is a conditional probability distribution of occurrences of context items, given the occurrence of the target. For each context item (c_j), the number of co-occurrences of the target item with it is calculated and divided by the total number of co-occurrences of w with each context item (see Eq. 2).

$$P(c_i|w) = \frac{f(c_i, w)}{\sum_{j=1}^n f(c_j, w)} \quad (2)$$

For each target item, the posterior probability distribution may be compared to the prior probability distribution of the context items. The prior gives the probabilities of encountering items without knowledge about the occurrence of the target and thus serves as a reference point: the more the posterior differs from it, the more information about its context a given target item brings. This information gain is exactly what relative entropy captures (see Eq. 3).

$$CD = D(P(c)||P(c|w)) = \sum_{i=1}^n P(c_i|w) \log_2 \frac{P(c_i|w)}{P(c_i)} \quad (3)$$

Results and discussion

CD values were calculated as outlined in the previous section for 33 inflections, separately in the child's sample and in the mother's sample. There were 4,610 noun tokens in the child's sample and 5,745 tokens in the mother's sample after all the exclusions. To avoid crediting the mother with more diversity, we took off 1,135 tokens from the latter at random. We repeated this 1,000 times, each time calculating CD for the inflections and comparing the mean difference between the mother

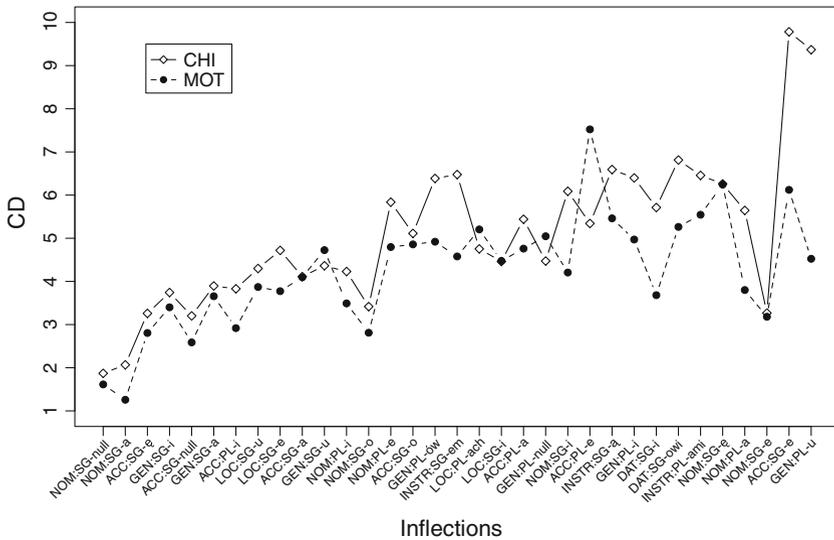


Fig. 3 CD values for noun inflections (endings marking case-number categories) in the child's sample and in the mother's sample (averaged across 1,000 samples)

and the child with a paired-sample t test. Here we report averaged results (Fig. 3 plots the average CD values and it can be seen that the child scores consistently higher than the mother for almost all inflections). The mother's mean was on average 4.248 ($SD = 0.047$), compared to the child's 5.081, with the standard deviation of the difference 1.26 ($SD = 0.08$). The average $t(32) = 3.81$ ($SD = 0.34$), $p < 0.002$ ($SD = 0.003$), and Cohen's $d = 0.66$ ($SD = 0.06$).

The observed difference between the child and her mother is highly significant and reflects a large effect size. Noun inflections are more informative about their immediate context, i.e., are more contextually distinct, in the child's sample than in the mother's sample. Importantly, this effect is evident even though the comparisons were based on the same set of inflections in both speakers, only when they were used with a given noun type by both speakers, and on the same number of tokens in both speakers' samples, and with the set of context words also controlled across the two speakers. We conclude that the immediate context in which noun inflections appear in the child's speech is highly restricted as compared to their context in the mother's speech, and that this can not be accounted for by differences in vocabulary size, utterance length, or sample size. This fact suggests that the observed difference results from the child's limited knowledge about possible contexts of use of noun inflections.

Two objections might be raised at this point. One is that if a noun occurs at the edge of an utterance, the co-occurrence window, initially set to ± 1 , shrinks to $+1$ or -1 only. Because the child's utterances were on average shorter than the mother's, a higher proportion of noun tokens in the child's sample are likely to involve such a reduced co-occurrence window, which might affect the difference between the

speakers.⁷ Therefore, we ran two additional analyses, explicitly setting the window to +1 or -1 for both speakers. Interestingly, while the difference for the preceding context only (-1) indeed decreased, the difference for the following context (+1) considerably increased. Overall, the difference between the two speakers remained significant.

The second objection might be that setting the window to adjacent items only, the purpose of which was not to credit the mother with more contextual diversity simply because she is more likely to produce longer utterances, might in some cases work against this control. For example, the child might know that a particular verb requires nouns in a particular inflectional form. If both the child and the mother use this verb with some nouns, but the mother occasionally inserts an adjective between them, with the window set to ± 1 her use will appear more diverse, even though it would not be caused by her knowing more contexts of use of this particular inflection.⁸ Indeed, a computational approach like the one we assumed in this study, although helps to avoid committing oneself to certain arbitrary decisions, brings about its own inaccuracies (measurement errors). On the other hand, this issue is in a way a vocabulary size problem. As in the example, in order to include this particular adjective as a context item, we would have to find it in the very same inflectional form in the child's sample as well, and this would be most likely if the child used it in the same or at least a similar construction as the mother. Hence, although we cannot directly control such artefacts, our rigorous control of the vocabulary size reduces their likelihood considerably.

General discussion

We have shown in this paper three things: (1) that a Polish two-year old child can use all case categories in both numbers, with virtually all endings possible, and making a highly limited number of morphological errors (i.e., supplying a wrong ending to a given stem); (2) that her use of those inflections, as compared to an adult, is in fact restricted in terms of the noun stems they combine with; and (3) that it is also restricted, compared to the adult, in terms of contexts they occur in. What makes our findings particularly sound are: (a) a densely and systematically sampled dataset, the first corpus of child Polish of this quality available in the field, (b) sophisticated quantitative methods, involving using adult input as a baseline for analyses and controlling for potentially biasing factors to ensure fair comparison. Let us elaborate briefly on the three findings before we discuss their theoretical implications.

As for the first two, our data support previous findings on Polish, with regard to low overgeneralisation rates in children's spontaneous use of noun inflections (Dąbrowska 2001, 2004; Smoczyńska 1985), and are in line with naturalistic studies on other languages, with regard to the early use of inflections being lexically restricted (Aguado-Orea 2004; Mueller Gathercole et al. 2002; Pizzuto and Caselli

⁷ We thank the anonymous reviewer for this suggestion.

⁸ We thank the Guest Editor for this suggestion.

1992; Rubino and Pine 1998). The latter finding offers a plausible explanation for the former, since it indicates that the child was rather conservative when combining stems and endings. If it is a typical pattern for Polish children, then it indicates that the early and simultaneous emergence of various endings is an effect of item-based learning, rather than a result of a precocious development of the inflectional system. Viewed from this perspective, the fact that children at the age of two may use noun forms in all 13 case-number combinations and with up to six different endings within a combination does not look so surprising any more. As our sample shows, two-year old children may have heard all those inflections numerous times. Even one of the least frequent inflections in the mother's sample of our dataset, the dative singular *-u*, occurred three times with two noun types, i.e., once with one type and twice with the other, which means—if we take these numbers as sample estimates—the child might have heard these noun types in this particular inflection 112 and 228 times respectively during her second year of life alone (assuming the corpus constitutes a 7% sample of the child's verbal activity over the 6-week recording period, a year consists of 52 weeks, and rounding all numbers down), which should make it possible for her to memorise the noun forms.

Other inflections appeared with more noun types than the dative singular *-u* did; some of them considerably more, even in this short 6-week recording period. It thus seems unlikely that all the noun productions from the child's sample had been learned by rote. Some limited productivity must be at work, as is also suggested by the few overgeneralisation errors.

With regard to the results of Study 3, this is a very interesting finding, since it shows—for the first time—that a child's use of inflections can be contextually restricted, just as it is lexically limited. Indeed, our results suggest that it is even more limited contextually than lexically, given a small effect size of the latter. However, a potential problem arises when considering Studies 2 and 3 together, since the latter might be thought of as yet another biasing factor for the former. In other words, the child may be using a limited number of inflections she knows with nouns she knows not because she cannot produce certain combinations (forms), but rather because she lacks a range of linguistic contexts to use them in. However, it could also be the other way around, i.e., the child's contextually restricted use of inflections may result from her limited ability to combine noun stems with endings; if she was more productive morphologically, she might be able to use inflections in a wider range of contexts. In the naturalistic setting it is impossible to identify the causative relation and clearly experimentation is needed to resolve this. Our results simply reveal this twofold nature of the restrictiveness of a young child's use of inflections. From a usage-based point of view this is not a problem, since different noun forms are heard by a child as parts of larger constructions and the process of abstracting inflections consists of these two intrinsically related aspects: abstracting across the stems that a given inflection applies to, and abstracting across the contexts in which it appears.

So how do our results fit into a broader theoretical context? In principle, they are not incompatible with the rule-based approach. For example, rules might emerge later in the course of development and a longitudinal study is clearly needed to reveal the full developmental path. On the other hand, the fact that, using a dense

sample, we found evidence of some, albeit not unrestricted, productivity at one point in time strongly suggests that the whole route, from no productivity to full productivity, is gradual rather than instantaneous. Furthermore, the focus on the English past tense in the rules vs. schemas debate in the child language field has led to the dual mechanism models gaining exclusive prominence. These models, emphasising the distinction between a default rule and memory-based irregulars, might be less adequate to account for systems other than the English past tense (Dąbrowska 2004; Dressler et al. 1996; Ravid et al. 2008, point out that even proponents of the dual mechanism approach already admit that they cannot account for the development of Dutch plurals). A thorough evaluation of the data from a rule-based point of view must thus await a formulation of a developmental model, which would deal specifically with a complex system like Polish noun inflections. The issues such a model would need to tackle include: (1) the distinction between regular rules and irregular forms, (1a) how the system can be divided between the two, and (1b) how the two can be combined, when switching between inflections (Krajewski et al. in press), (2) the scope of the application of rules, (2a) whether partially productive (micro-)rules are possible (Dressler 1999), and (2b) what the set of possible factors conditioning rules is (e.g., abstract gender classes, phonological features, semantics), (3) the potential role of context of use: how the context requiring a given rule is incorporated into its formulation (e.g., in terms of fully abstract symbolic categories based on conventional case distinctions, like the genitive, the accusative and so on). The model would need to consider all the above points and, at the same time, it would need to be developmentally sound, i.e., it would have to be explicit as to how rules emerge/are constructed/formulated, and it would have to be able to account for the transition from the use of rote-learned unanalysed forms to the productive application of rules. Moreover, it would have to be able to account for available data, e.g., the role of frequency and similarity found in various experimental studies (Dąbrowska 2006; Dąbrowska and Szczerbiński 2006; Krajewski et al. in press), and differences in the productive use across different inflections in adults (Dąbrowska 2004, 2008).

As for the schema-based approach, it is not only compatible with the data, but also the results are exactly as this approach would predict: lexically restricted use, accounting for a low rate of overgeneralisation errors, can be explained by slow and gradual abstraction of the form specification of a schema, and contextually restricted use of inflections can be explained by an analogical slow and gradual abstraction of the meaning specification of a schema. However, this approach also needs a specific model accounting for an inflectional system like Polish noun inflections. Such a model would have to account for the contextual restrictiveness and explain its relationship with the lexical restrictiveness. It would also need to be more explicit about the role of frequency and similarity in abstracting schemas: (1) what units they are based on (phonemes, syllables, morphemes etc.), and whether the units are constant or change dynamically (becoming more abstract in the course of development), (2) how the two factors are defined and calculated (e.g., given that the productive use of inflections involves switching between various pairs of them, how the relative frequency should be calculated), and (3) the relative importance of, and relationship between, frequency and similarity.

Acknowledgments The authors would like to thank Ela Łazarewicz-Wyrzykowska for her help with coding reliabilities, and Ewa Dąbrowska, the anonymous reviewer, and the Editor of the Special Issue for their very helpful and constructive feedback.

References

- Abbot-Smith, K., & Behrens, H. (2006). How known constructions influence the acquisition of other constructions: The German passive and future constructions. *Cognitive Science*, 30(6), 995–1026.
- Aguado-Orea, J. (2004). *The acquisition of morpho-syntax in Spanish: Implications for current theories of development*. Unpublished PhD, The University of Nottingham.
- Berman, R. A. (2004). Between emergence and mastery: The long developmental route of language acquisition. In R. A. Berman (Ed.), *Language development across childhood and adolescence* (pp. 9–34). Amsterdam: John Benjamins.
- Burgess, C., & Lund, K. (1997). Modelling parsing constraints with high-dimensional context space. *Language and Cognitive Processes*, 12(2/3), 177–210.
- Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.
- Bybee, J. L. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10(5), 425–455.
- Cameron-Faulkner, T., Lieven, E. V. M., & Theakston, A. L. (2007). What part of *no* do children not understand? A usage-based account of multiword negation. *Journal of Child Language*, 34(2), 251–282.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Dąbrowska, E. (1997). *Cognitive semantics and the Polish dative*. Berlin: Mouton de Gruyter.
- Dąbrowska, E. (2001). Learning a morphological system without a default: The Polish genitive. *Journal of Child Language*, 28(3), 545–574.
- Dąbrowska, E. (2004). Rules or schemas? Evidence from Polish. *Language and Cognitive Processes*, 19(2), 225–271.
- Dąbrowska, E. (2005). Productivity and beyond: Mastering the Polish genitive inflection. *Journal of Child Language*, 32(1), 191–205.
- Dąbrowska, E. (2006). Low-level schemas or general rules? The role of diminutives in the acquisition of Polish case inflections. *Language Sciences*, 28(1), 120–135.
- Dąbrowska, E. (2008). The effects of frequency and neighbourhood density on adult speakers' productivity with Polish case inflections: An empirical test of usage-based approaches to morphology. *Journal of Memory and Language*, 58(4), 931–951.
- Dąbrowska, E., & Lieven, E. V. M. (2005). Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics*, 16(3), 437–474.
- Dąbrowska, E., & Szczerbiński, M. (2006). Polish children's productivity with case marking: The role of regularity, type frequency, and phonological diversity. *Journal of Child Language*, 33(3), 559–597.
- Dressler, W. U. (1999). Why collapse morphological concepts? *Behavioral and Brain Sciences*, 22(6), 1021–1021.
- Dressler, W. U., Drażyk, R., Drażyk, D., Dziubalska-Kołaczyk, K., & Jagła, E. (1996). On the earliest stages of acquisition of Polish declension. In C. Koster & F. Wijnen (Eds.), *Proceedings of the Groningen Assembly on language acquisition* (pp. 185–195). Groningen: Centre for Language and Cognition.
- Dressler, W. U., Dziubalska-Kołaczyk, K., & Fabiszak, M. (1997). Polish inflection classes within natural morphology. *Bulletin de la Société Polonaise de Linguistique*, 53, 95–119.
- Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: SAGE.
- Goldberg, A. E. (2005). *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.
- Howell, D. C. (2007). *Statistical methods for psychology* (6th ed.). Belmont, CA: Thomson/Wadsworth.
- Kirjavainen, M., Theakston, A. L., & Lieven, E. V. M. (2009a). Can input explain children's me-for-I errors? *Journal of Child Language*, 36(5), 1091–1114.
- Kirjavainen, M., Theakston, A. L., Lieven, E. V. M., & Tomasello, M. (2009b). 'I want hold Postman Pat': An investigation into the acquisition of infinitival marker 'to'. *First Language*, 29(3), 313–339.

- Köpcke, K.-M. (1998). The acquisition of plural marking in English and German revisited: schemata versus rules. *Journal of Child Language*, 25(2), 293–319.
- Kovačević, M., Palmović, M., & Hržica, G. (2009). The acquisition of case, number and gender in Croatian. In U. Stephany & M. D. Voeikova (Eds.), *Development of nominal inflection in first language acquisition: A cross-linguistic perspective* (pp. 153–177). Berlin: Mouton de Gruyter.
- Krajewski, G. (2005). The role of grammatical gender in the acquisition of noun inflection in Polish. *Psychology of Language and Communication*, 9(1).
- Krajewski, G., Theakston, A. L., Lieven, E. V. M., & Tomasello, M. (in press). How Polish children switch from one case to another when using novel nouns: Challenges for current models of inflectional morphology. *Language and Cognitive Processes*.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211–240.
- Lieven, E. V. M., Behrens, H., Speares, J., & Tomasello, M. (2003). Early syntactic creativity: a usage-based approach. *Journal of Child Language*, 30(2), 333–367.
- MacKay, D. J. C. (2003). *Information theory, inference and learning algorithms*. Cambridge: Cambridge University Press.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Marchman, V. A. (1997). Children's productivity in the English past tense: The role of frequency, phonology, and neighborhood structure. *Cognitive Science*, 21(3), 283–304.
- Maslen, R. J., Theakston, A. L., Lieven, E. V. M., & Tomasello, M. (2004). A dense corpus study of past tense and plural overregularization in English. *Journal of Speech, Language, and Hearing Research*, 47(6), 1319–1333.
- Matthews, D. E., & Theakston, A. L. (2006). Errors of omission in English-speaking children's production of plurals and the past tense: The effects of frequency, phonology, and competition. *Cognitive Science*, 30(6), 1027–1052.
- McDonald, S. A., & Shillcock, R. C. (2001). Rethinking the word frequency effect: The neglected role of distributional information in lexical processing. *Language and Speech*, 44(3), 295–323.
- Mueller Gathercole, V. C., Sebastián, E., & Soto, P. (2002). The emergence of linguistic person in Spanish-speaking children. *Language Learning*, 52(4), 679–722.
- Pine, J. M., Lieven, E. V. M., & Rowland, C. F. (1998). Comparing different models of the development of the English verb category. *Linguistics*, 36(4), 807–830.
- Pinker, S. (1997). *How the mind works*. London: Penguin.
- Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, 6(11), 456–463.
- Pizzuto, E., & Caselli, M. C. (1992). The acquisition of Italian morphology: Implications for models of language development. *Journal of Child Language*, 19(3), 491–557.
- Ravid, D. (2004). Derivational morphology revisited: Later lexical development in Hebrew. In R. A. Berman (Ed.), *Language development across childhood and adolescence* (pp. 53–82). Amsterdam: John Benjamins.
- Ravid, D., Dressler, W. U., Nir-Sagiv, B., Korecky-Kröll, K., Souman, A., Rehfeldt, K., Laaha, S., Bertl, J., Basboell, H., & Gillis, S. (2008). Core morphology in child directed speech: Crosslinguistic corpus analyses of noun plurals. In H. Behrens (Ed.), *Corpora in language acquisition research: History, methods, perspectives* (pp. 25–60). Amsterdam: John Benjamins.
- Ravid, D., & Farah, R. (2009). Noun plurals in early Palestinian Arabic: a longitudinal case study. In U. Stephany & M. D. Voeikova (Eds.), *Development of nominal inflection in first language acquisition: a cross-linguistic perspective* (pp. 411–432). Berlin: Mouton de Gruyter.
- Rubino, R. B., & Pine, J. M. (1998). Subject-verb agreement in Brazilian Portuguese: What low error rates hide. *Journal of Child Language*, 25(1), 35–59.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell Systems Technical Journal*, 27, 379–423.
- Smoczyńska, M. (1985). The acquisition of Polish. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 1: The data, pp. 595–686). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Stephany, U., & Voeikova, M. D. (Eds.). (2009). *Development of nominal inflection in first language acquisition: A cross-linguistic perspective*. Berlin: Mouton de Gruyter.
- Theakston, A. L., & Lieven, E. V. M. (2008). The influence of discourse context on children's provision of auxiliary BE. *Journal of Child Language*, 35(1), 129–158.

-
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, M., & Stahl, D. (2004). Sampling children's spontaneous speech: How much is enough? *Journal of Child Language*, *31*(1), 101–121.
- Vargha, A., & Delaney, H. D. (2000). A critique and improvement of the CL Common Language effect size statistics of McGraw and Wong. *Journal of Educational and Behavioral Statistics*, *25*(2), 101–132.