J. Child Lang. **33** (2006), 559–597. © 2006 Cambridge University Press
doi:10.1017/S0305000906007471 Printed in the United Kingdom

Polish children's productivity with case marking: the role of regularity, type frequency, and phonological diversity*

EWA DĄBROWSKA AND MARCIN SZCZERBIŃSKI
University of Sheffield

(*Received* 3 *March* 2004. *Revised* 30 *August* 2005)

ABSTRACT

57 Polish-speaking children aged from 2;4, to 4;8 and 16 adult controls participated in a nonce-word inflection experiment testing their ability to use the genitive, dative and accusative inflections productively. Results show that this ability develops early: the majority of two-year-olds were already productive with all inflections apart from dative neuter; and the overall performance of the four-year-olds was very similar to that of adults. All age groups were more productive with inflections that apply to large and/or phonologically diverse classes, although class size and token frequency appeared to be more important for younger children (two- and three-year-olds) and phonological diversity for older children and adults. Regularity, on the other hand, was a very poor predictor of productivity. The results support usage-based models of language acquisition and are problematic for the dual mechanism model.

INTRODUCTION

One of the most controversial issues in the recent literature on language development concerns the nature of the mental mechanisms underlying linguistic creativity, that is to say, our ability to produce previously unheard combinations of words and morphemes. Because morphological productivity, or the ability to supply inflected forms such as the plural of a noun or the

[*] We would like to thank Ewa Borek, Mariola Busławska, Małgorzata Ciołek, Ewa Czerlińska, Celina Kośmider, Małgorzata Michalak, and Bożena Pławska for their help in collecting the data, Barbara Dąbrowska for organizational support throughout the duration of the project, and Grzegorz Krajewski for his assistance with coding the corpus data. Very special and warm thanks go to the children from Żłobek nr 1 and Przedszkole nr 81 in Gdańsk who participated in the experiment. This study was supported by British Academy grant RB 100556 awarded to the first author. Address for correspondence: Dr. Ewa Dąbrowska, Department of English Language and Linguistics, University of Sheffield, Sheffield S10 2TN, UK. e-mail: e.dabrowska@shef.ac.uk

past tense of a verb, is relatively easy to study, much of the research addressing this issue has focused on this particular aspect of our creative abilities.

There are two main approaches to explaining morphological productivity. According to the dual mechanism theory (Marcus *et al*., 1992, 1995; Pinker, 1998, 1999; Clahsen, 1999; Clahsen *et al*., 2002), the mainspring of our ability to produce novel forms is a mental device which implements symbolic rules of the kind traditionally proposed by linguists, e.g. 'to form the past tense of a regular English verb, add *-ed* to the verb stem'. In addition to this, human beings are also equipped with a second, relatively 'low-tech' mechanism, associative memory, which is used for storing and retrieving irregular forms. Irregular forms, though not fully predictable, tend to share certain similarities, which are captured by means of partially overlapping memory representations. This reinforces similarities, allowing the memory system to extract recurring patterns which can then be applied to other words. Thus, the memory system is also productive. Crucially, however, its productivity is fairly limited: it is constrained by phonological similarity (an associative network can generalize a pattern to a new word only if it resembles a previously learned word), and it is also strongly dependent on frequency (only relatively frequent patterns are generalized to novel items). Symbolic rules, on the other hand, can be used with any stem, regardless of its phonological properties, and act as a default system which must apply when the speaker's lexicon lacks an entry for a particular word (i.e. with newly coined words, recent borrowings, etc.) and when the information stored in memory is inaccessible (for example, when the memory trace is too weak, or when information stored with the root cannot be passed on to the derived form). Marcus *et al*. (1995) list sixteen such circumstances, and propose that children identify the default inflection by noting that it is used in 'one or two' of them and apply it across the board to all words except those which are already marked as irregular in their mental lexicons. It follows, then, that the most important predictor of productivity is regularity (in the technical sense of the theory, i.e. default status), not frequency: the regular pattern can in fact apply to a relatively small number of items (Marcus *et al*., 1995).

The alternative position, associated with cognitive linguistics (Langacker, 1991, 2000; Bybee, 1995; Taylor, 2002; Dąbrowska, 2004), most connectionist models (e.g. Rumelhart & McClelland, 1986; Plunkett & Marchman, 1993; Elman *et al*., 1996), and other 'usage-based' theories, is that a single mental mechanism – schemas of varying degrees of generality – can account for our ability to supply both regular and irregular inflections. In the course of language acquisition, learners memorize many inflected forms and note that they share certain properties. The commonalities are reinforced through repeated use, which results in the extraction of schemas.

The original function of schemas is to capture redundancies in the lexicon; however, once they become well-established, they can be used to inflect novel words. Usage-based theories emphasize the importance of frequency, especially type frequency, and see learners as basically conservative or 'lazy' in that they naturally prefer low-level generalizations over clusters of phonologically similar forms to more abstract rules which apply 'across the board'.

The Polish case marking system

Much of the recent research on the development of morphological productivity has focused on the English past tense, which confounds regularity, type frequency, and phonological diversity: the regular inflection applies to a much larger number of verbs than any of the irregular patterns, and it is also the only pattern which is phonologically unrestricted (the irregular patterns apply to individual verbs or to clusters of verbs sharing certain phonological properties). Furthermore, regular and irregular verbs in English rely on a different morphological mechanism to form the past tense: the past tense of regular verbs is formed by suffixation (e.g. *walk-ed*), while most irregular verbs require stem changes (*sit – sat, catch – caught, sing – sang*, etc.). Because of these confounding factors, it is impossible to determine whether any observed differences between regular and irregular inflections are attributable to regularity as such or to frequency, phonological diversity, morphological mechanism, or some combination of these factors. It is clear, therefore, that in order to be able to evaluate the approaches outlined above, we will need to consider evidence from other languages.

In this study, we present data on the acquisition of parts of the Polish case-making system, specifically, the suffixes marking the genitive, dative and accusative singular. Polish is a morphologically rich language with a fairly elaborate system of case inflections. There are seven cases, each signalled by several different suffixes (*cf.* Table 1); their main uses are summarized in Table 2. All case endings are portmanteau morphs signalling number as well as case. In this paper, we will confine ourselves to the nine affixes printed in boldface in Table 1.

The single most important determinant of the choice of ending is gender, which can be fairly reliably predicted from the phonological form of the nominative: feminine nouns typically end in *-a*, masculines in a consonant, and neuters in *-o, -e*, or *-ę*.[1] When a particular case has more than one ending for a given gender, other factors come into play. The choice of the

[1] Nouns with non-canonical phonological properties (i.e., masculines which do not end in a consonant, feminines which do not end in *-a*, and neuters which do not end in *-o, -e*, or *-ę*) account for about 2% of the nouns in the input in the Marysia corpus, and about 10% of the nouns found in a fairly large contemporary dictionary (Bańko, 2000).

TABLE 1. *The Polish case marking system* (*the singular endings*)

Case	Feminine	Masculine	Neuter
Nominative	*-a* (*-Ø*, *-i*)	*-Ø* (*-a*, *-o*)	*-o*, *-e*, *-ę*
Genitive	***-i*/*-y***	***-a***, *-u* (*-i*/*-y*)	***-a***
Dative	***-ʹe***, *-i*/*-y*	***-owi*** (*-u*, *-ʹe*, *-i*/*-y*)	***-u***
Accusative	***-ę*** (*-Ø*)	*-Ø*, ***-a*** (*-ę*, *-o*)	**=NOM**
Instrumental	*-ą*	*-em* (*-ą*)	*-em*
Locative	*-ʹe*, *-i*/*-y*	*-ʹe*, *-u* (*-i*/*-y*)	*-ʹe*, *-u*
Vocative	*-o*, *-u*, *-i*/*-y*, (*-Ø*)	*-ʹe*, *-u*, (*-o*)	*-o*, *-e*, *-ę*

Note: Endings in parentheses are restricted to fairly narrow classes of exceptions. The [ʹ] symbol before an ending indicates that it triggers palatalization of the preceding consonant(s). The distribution of *-i* and *-y* is governed by very general phonotactic constraints and they are therefore regarded as variants of the same ending.

TABLE 2. *Polish cases and their main uses*

Case	Main uses
Nominative	citation form; subject; subject predicative; when addressing a person
Genitive	adnominal modifier (e.g. possessor); partitive; after negated verbs and after *nie ma* 'not have:3SG' to indicate absence; with certain verbs and prepositions
Dative	indirect object (addressee, recipient); beneficiary; experiencer; with certain verbs and prepositions
Accusative	direct object (patient, theme); with certain prepositions
Instrumental	instrument; subject predicative; with certain verbs and prepositions
Locative	with certain prepositions
Vocative	when addressing a person; sometimes as subject

feminine dative ending is determined by phonological factors. The so-called 'soft' stems (those ending in [l] or a consonant with the feature [+PALATAL]) take *-i*; 'hardened' stems (those ending in an unpalatalized affricate or a post-alveolar fricative) take *-y*; and 'hard' stems (ending in any other consonant, i.e. a labial, dental, or velar fricative; a stop; an unpalatalized nasal, [w], or [r]) take *-ʹe*. Addition of the *-ʹe* ending triggers obligatory changes in the final consonant or consonants of the stem (see Tokarski, 2001 for details). The distribution of the accusative masculine endings is determined primarily by semantic factors: animate nouns take *-a* and inanimate nouns normally take *-Ø*, although the *-a* ending has spread to some inanimate nouns. The choice of the genitive masculine ending also depends to some extent on semantics (animate nouns and most nouns designating small easily manipulable objects take *-a*; nouns designating substances, collections of objects, and abstract concepts usually take *-u*); but morphological and phonological criteria are also relevant (see Westfal, 1956; Bodnarowska, 1962; Kottum, 1981). However, apart from animacy,

TABLE 3. *Endings used in the 'default circumstances' identified by Marcus et al.* (1995)

Circumstance	GEN FEM *-i/-y*	GEN NEU *-u*	DAT MAS *-owi*	DAT FEM *- e*	DAT NEU *-u*	ACC MAS *-a*	ACC FEM *-ę*
Low-frequency words	yes	yes	yes	yes	yes	yes	yes
Unusual-sounding words	no	no	no	no	no	no	no
Onomatopoeia	N/A	no	N/A	N/A	no	N/A	N/A
Word is mentioned rather than used	yes	yes	yes	yes	yes	yes	yes
Surnames	yes	N/A	yes	yes	N/A	yes	yes
Borrowings	yes	yes	yes	yes	yes	yes	yes
Truncations	yes	no	yes	yes	no	yes	yes
Acronyms	no	no	no	no	no	no	no
Derivation from a different category:							
(a) affixation	yes	yes	yes	yes	yes	yes	yes
(b) backformation	yes	yes	yes	yes	yes	yes	yes
(c) nominalized adjectives	no	no	no	no	no	no	no
Derivation via different category	yes	yes	yes	yes	yes	yes	yes
Derivation via name	yes	N/A	yes	yes	N/A	yes	yes
Bahuvrihi compounds and nominalized phrases	yes	yes	yes	yes	yes	yes	yes

none of the criteria are very reliable, and they are often in conflict, so genitive endings – at least the genitive endings of inanimate nouns – must be learned on an item-by-item basis. Thus, the genitive masculine inflection is irregular in the sense of being largely unpredictable. It is also irregular in the technical sense of the dual mechanism theory: that is to say, neither of the two endings is uniquely associated with the circumstances which, according to Marcus *et al*. (1995), call for the default inflection, and hence neither can be regarded as the default (for a detailed discussion, see Dąbrowska, 2001, 2004).

The other non-parenthesized genitive, dative and accusative endings given in Table 1, on the other hand, must be regarded as regular. They are either the only or the most widely applicable endings available for nouns belonging to the relevant subclass, and consequently apply to the vast majority of nouns in their domain of application, while the alternatives given in parentheses apply to narrowly defined classes of exceptions. Furthermore, they are used in most, though not all, of the 'default circumstances' identified by Marcus *et al*. (1995).[2] For the endings used in the experiment, this is demonstrated in Table 3; for further discussion, see Dąbrowska, 2004.

[2] There are some systematic exceptions (unusual-sounding words and acronyms are normally not inflected at all; nominalized adjectives take adjectival rather then nominal endings), as well as some circumstances in which the criteria are not applicable (surnames cannot be neuter; renderings of sounds are always neuter). Two other circumstances from the Marcus *et al*. list (homophones and rhymes) are not listed in Table 3

TABLE 4. *The main markers of the genitive, dative and accusative case and the number of nouns with each ending in the input to a two-year-old child*

Case	Ending	Tokens	Types
Genitive	masculine *-a*	64	40
	masculine *-u*	35	18
	feminine *-i/-y*	94	44
	neuter *-a*	26	19
	other	1	1
	total	**220**	**122**
Dative	masculine *-owi*	14	7
	feminine *- e*	9	5
	feminine *-i/-y*	16	6
	neuter *-u*	1	1
	other	4	3
	total	**44**	**23**
Accusative	masculine *-a*	30	20
	masculine *-Ø*	108	51
	feminine *-ę*	158	79
	neuter = NOM	47	14
	other	3	2
	total	**346**	**166**

Case inflections also differ along dimensions other than regularity. The masculine and feminine classes are relatively large, so most masculine and feminine inflections apply to more nouns than the neuter ending. However, in the dative, the feminine class is split between two endings, and in the genitive and the accusative, there are two masculine endings, so the domains of application of the endings for each gender vary in size, depending on the case. Table 4 gives information about the size of the domains of application of the principal markers of the three cases that are the object of this study, calculated on the basis of a five-hour sample of speech addressed to a two-year-old girl (the Marysia corpus: see Dąbrowska, 2005).

The classes of nouns to which the various case inflections apply also differ in phonological diversity, which has been argued to affect productivity (Dąbrowska, 2004). There is no standard way of quantifying diversity. However, since the choice of inflectional ending often depends on properties of the final syllable, and since it is known that children pay attention to the ends of words (*cf.* Slobin, 1985), the most useful measures would be those

because they are not useful for identifying regular inflections (the fact that homophones and rhymes take different endings does not tell us *which* of the endings is regular).

Note that the fact that the circumstances in which regular endings are used do not match up crosslinguistically is problematic for the dual mechanism theory. Even more difficult to accommodate within the theory is the fact that the genitive masculine endings *-a* and *-u*, which are undeniably irregular, also occur in many 'default' circumstances (see Dąbrowska, 2004 for a detailed discussion).

TABLE 5. *Phonological diversity of the domain of application of the case affixes used in the experiment*

Ending	Attested diversity	Potential diversity
genitive masculine *-a*	32	50
genitive feminine *-i/-y*	19	28
genitive neuter *-a*	7	9
dative masculine *-owi*	7	82
dative feminine *- e*	2	14
dative neuter *-u*	1	9
accusative masculine *-a*	19	29
accusative feminine *-ę*	21	24
accusative neuter = NOM	6	9

based on the amount of variation found in the final syllable. For the purposes of this study, we used two measures based on the number of distinct final syllables found in words belonging to a particular class: POTENTIAL diversity and ATTESTED diversity. Both measures were computed on the basis of nouns which occurred in adult utterances in the first five files of the Marysia corpus. To compute potential diversity, we extracted all the nouns which occurred in the genitive, dative, or accusative singular and divided them into classes according to which ending they require in a particular case (e.g. nouns which take the dative masculine *-owi*, nouns which take the dative feminine *- e*, and so on). We then counted (for each class separately), the number of different final syllables in the citation form (i.e. the nominative).[3] Attested diversity was computed by extracting the nouns which were used with a particular ending in the corpus (e.g. nouns which actually occurred with the dative masculine *-owi*, etc.) and then counting the number of distinct final syllables in the citation forms of these nouns. This last measure reflects not just the diversity found in nouns belonging to a particular inflectional subclass, but also frequency (since high-frequency forms were more likely to be captured in a particular form in the sample). Potential diversity, on the other hand, is a purer measure of the amount of diversity associated with a particular inflectional class in the language.

Table 5 shows the values of the two measures of phonological diversity for the nine endings that were elicited in our study. As can be seen from the table, the figures for potential diversity are the highest for masculine endings

[3] The implicit assumption behind this measure is that children learn most nouns in the nominative, and that productive use of case morphology involves converting the nominative into the oblique form required by the grammatical context in which the noun occurs. There is good evidence that most nouns are indeed learned in the nominative: it is, by a wide margin, the most frequent case in the input to young children (*cf.* Table 7); and young children often use the nominative in grammatical contexts which require oblique cases, while the opposite kind or error is extremely rare.

TABLE 6. *Correlations between frequency and phonological heterogeneity* (*Spearman's rho,* N = 9)

	type frequency	attested diversity	potential diversity
token frequency	0·950***	0·824**	0·271
type frequency		0·908***	0·356
attested diversity			0·633†

*** Correlation is significant at the 0·001 level (2-tailed).
** Correlation is significant at the 0·01 level (2-tailed).
† Correlation approaches significance ($p < 0{\cdot}10$, 2-tailed).

and the lowest for neuter endings, with feminine endings falling in between.[4] For attested diversity, the picture is rather similar, except in the accusative, where the figure for the feminine *-ę* is slightly higher than for the masculine *-a*; this is due to the fact that there are two different masculine accusative endings, and *-a* actually applies to a relatively small set of nouns whose members inevitably have fewer distinct final syllables (see Table 4).

As can be seen from Table 6, the three corpus-based measures (type frequency, token frequency, and attested diversity) are strongly correlated. Potential diversity is moderately strongly correlated with actual diversity, and relatively independent of the other measures. This confirms the intuition that attested diversity is also an indirect measure of type frequency.[5]

Finally, the cases themselves differ in frequency. Table 7 shows the frequencies (converted into percentages) of all singular case forms in the language addressed to Marysia. As we can see from the table, nominatives comprise over a half all the noun tokens in the input. Genitive and accusative forms are also very frequent, accounting for about 12% and 19% respectively of the noun tokens. The dative case, however, is quite rare: only about 2% of the noun tokens in the corpus carry one of the dative inflections.[6]

[4] The fact that masculine endings apply to phonologically more diverse classes is due to the fact that they nearly always end in a consonant, and hence have a variety of offsets. Feminine and neuter nouns, on the other hand, typically end in *-a* and *-o* respectively, and thus cluster in regions of phonological space defined by the final open syllable (*-ta, -ka, -to, -ko*, etc.).

[5] The *p* value for the correlation between attested and potential diversity for the nine endings used in the study is 0·068, and thus only approaches the conventional $p < 0{\cdot}05$ level of significance. Note, however, that the correlations were computed only for the nine endings elicited in the experiment. If we include data for the other productive endings listed in Table 4 (masculine genitive *-u*, feminine dative *-i*/*-y*, and accusative masculine *-Ø*), the correlation is clearly significant (rho = 0·675, $p = 0{\cdot}016$).

[6] This is partly due to the fact that NPs expressing dative roles (experiencers, recipients, addressees) tend to be highly topical, and hence are often realized as pronouns. As a result, children get considerably more experience of the dative as a syntactic category than the figures in the table suggest.

TABLE 7. *Frequency of singular noun forms in the input to a two-year-old child*

Case	Percent in input ($N = 1848$)
Nominative	54
Genitive	12
Dative	2
Accusative	19
Instrumental	4
Locative	4
Vocative	5

In summary, the various inflections differ in

1. regularity: the genitive masculine inflections are irregular; all other inflections (except those given in parentheses in table 1) can be regarded as regular;
2. the size of the domain of application: masculine and feminine inflections apply to more noun types than neuter inflections; the dative masculine inflection applies to more nouns than the dative feminine inflection;[7] and the accusative feminine inflection applies to more nouns than the accusative masculine inflection;
3. overall frequency: nouns with genitive and accusative marking are more frequent than nouns marked for the dative;
4. the phonological structure of the domain of application: masculine inflections apply to phonologically more diverse classes than feminine and neuter inflections.

Earlier research on the acquisition of Polish case inflection

Earlier research with spontaneous speech, summarized in Smoczyńska (1985), suggests very early mastery of the case marking system. All the major singular markers and the nominative and accusative plural emerge before age 2, and are used correctly from the very beginning: except in a few isolated areas, error rates are extremely low. Dąbrowska (2001, 2004) confirms these findings, but observes that error rates are also very low on masculine nouns occurring in genitive contexts, which suggests that children memorize many ready-made inflected forms as well as extracting rules

[7] The figures in Table 4 on which this prediction is based are quite small: the masculine *-owi* ending occurred with 14 tokens of 7 noun types and the feminine *-e* with only 9 tokens of 5 types. Therefore, we conduced a second tally of the frequency of the two endings, counting all the nouns in Marysia's productive vocabulary ($N = 1030$). This confirmed that *-owi*, which was required by 48% of the nouns, is indeed considerably more frequent than *-e*, which applied to only 38% of the nouns.

which capture recurrent patterns. She also shows that the highly irregular genitive masculine inflections and the regular feminine inflections develop in parallel; there is no evidence that the former is more difficult, or that children treat any one ending as the default. Both of these findings are problematic for the dual mechanism theory.

Predictions

In the following experiment, we tested children's ability to supply nine inflections: genitive masculine *-a*, genitive feminine *-i*/*-y*, and genitive neuter *-a*; dative masculine *-owi*, dative feminine - *e*, and dative neuter *-u*; and accusative masculine *-a*, accusative feminine *-ę*, and accusative neuter *-o*/*-ę*. We chose these inflections because they are relatively easy to elicit, and because the two theories discussed above make different predictions about the ease of their acquisition.

According to usage-based theories, one of the most important factors determining productivity is frequency, especially type frequency: patterns which apply to a large number of nouns should be easier to generalize, and thus acquired earlier. If this hypothesis is correct, children should perform better (i.e. supply the correct endings more frequently) on masculine and feminine inflections than on neuter inflections, since the masculine and feminine classes are considerably larger than the neuter class. We would also expect them to perform better on genitive and accusative inflections, both of which are very frequent, than on dative inflections, which are relatively rare in terms of token frequency and for the most part restricted to animate nouns. Usage-based theories also predict case × gender interactions: children should perform better on masculines than on feminines in the dative, and better on feminines than on masculines in the accusative. Finally, usage based models predict better performance on familiar words than on unfamiliar words. This is because the inflected forms of many familiar words can be retrieved from memory, and hence do not need to be computed.

According to the dual-mechanism theory, the most important predictor of productivity is regularity: regular inflections are vastly more productive than irregular ones. Thus, if the theory is correct, we should expect children to perform better on genitive feminine and genitive neuter inflections than on the genitive masculine (because the latter lacks a default). There should be no other differences between the gender classes (since regular inflections are not sensitive to frequency or phonological similarity).

The dual mechanism theory does not predict case and lexicality effects but can accommodate them. Poorer performance on the dative can be explained by extra-grammatical factors such as semantic complexity: the notion of experiencer is arguably more abstract than that of possessor or patient; and concepts such as recipient and addressee presuppose scenes

involving three participants (giver, gift, and recipient; speaker, message, and addressee). Lexicality effects can be accommodated by appealing to performance factors: inflecting a recently learned word may be more demanding than inflecting a familiar word.

METHOD

Participants

The participants were 57 children (29 boys and 28 girls) aged 2;4–2;8 (mean 2;7), 3;2–3;8 (mean 3;5) and 4;3–4;8 (mean 4;5) from a crèche and kindergarten in the Gdańsk area and 16 adult controls (mean age 38 years). The control group consisted of staff from the same institutions and two mothers – i.e. individuals who provided a substantial part of the children's input. All participants spoke Polish as their first language.

Materials

Linguistic stimuli. The stimuli were 12 real nouns and 12 nonce nouns. Within each of these categories, there were four words of each gender. Half of the real words of each gender were diminutives and half were simplex forms; among the nonce words, half could be analysed as diminutives (that is to say, they ended in *-ek*, *-ka*, and *-ko*, the most productive diminutive affixes), and half were non-diminutives.[8] All the words were phonotactically legal sequences of sounds and had gender-typical endings: the masculine words ended in a hard consonant, the feminine words ended in *-a*, and the neuters ended in *-o* or *-ę*. A full list of all the words used in the experiment is given in the Appendix.

The real words referred to familiar animals (cow, sheep, frog, etc.) The masculine and feminine real words were matched for frequency and were all high-frequency items that were almost certainly familiar to the two-year-olds. This was not possible with the neuter words. Most neuter nouns refer to inanimate objects, and because of the nature of the task (see below), the referents had to be animate, and therefore we were obliged to use words of lower frequency. Thus, while we can assume that virtually all of the children were familiar with the real masculine and feminine nouns, we cannot be sure that they knew all the neuters, and hence the results for real neuter nouns are not comparable with the results for real nouns of other genders. While this is not ideal, we decided that it was better than the alternative, which would have been to use less frequent masculines and feminines. Our main focus is the children's ability to inflect nonce words: the real words are

[8] The effect of morphological structure of the stem on children's productivity with case endings is the object of another study (Dąbrowska, 2006).

a control condition, included in the experiment in order to determine whether the children had understood the task and whether they knew which case was required in the grammatical context in which the test words were used. Accordingly, we decided that it was better to have good controls for two of the three genders which may not be comparable with controls for the third gender than to have uninterpretable controls for all three genders.

The stimuli were divided into four blocks, each containing six words (one nonce word of each gender and one real word of each gender; if the nonce words were simplex the real words were diminutive and vice versa). There were two versions of the experiment: nouns presented in the simplex form in version A were diminutive in version B and vice versa. Half of the participants were tested with version A and the other half with version B.

Other materials

24 toy animals (12 familiar and 12 unfamiliar) were used as referents for the test words. A further 8 toy animals were used in practice items, and 10 small objects were used as 'presents' for the animals (see the section on eliciting the dative inflection).

Procedure

The children were tested in a quiet room in their crèche or kindergarten by two experimenters. Experimenter 1 interacted with the child. Experimenter 2 prepared the props, kept a live log of the child's responses, and audio recorded the testing sessions for later checking. The experimenters were recruited from the teachers and carers employed in the day care centres which the children attended, and hence were well known by the children.

Each session consisted of three phases: presentation, recognition, and testing.

Presentation phase

Experimenter 1 showed the child an unfamiliar toy and introduced the corresponding nonce word in the citation form (the nominative) in a presentational construction. This was followed by a simple definition, e.g.

(1) *Zobacz, to jest gryma.*
'Look, this is a gryma.'
(2) *Gryma to jest taki różowy stworek.*
'A gryma is a pink creature like this one.'

The child was then asked to repeat the word:

(3) *Potrafisz powiedzieć* gryma?
'Can you say *gryma*?'

If the child did not succeed in imitating the word correctly, she was asked to try again. Then the experimenter repeated the word and demonstrated some action involving the referent, for example:

(4) *Tak*, gryma, *bardzo dobrze!*
'Yes, *gryma*, very good!'
(5) *Ta gryma lubi skakać.*
'This gryma likes to jump.'
(6) *Zobacz, jak gryma skacze, hop! hop! hop!* [demonstrates]
'Look how the gryma jumps, boing! boing! boing!'

Then the toy was handed over to the child, who was asked to perform the same action:

(7) *A teraz ty pokaż jak gryma skacze.*
'Now you show how the gryma jumps.'

The same procedure was repeated for the remaining two nonce words in the block. The experimenter then presented the familiar animal toys one at a time and asked the child if she knew what they were called. If the child did not supply the target word, the experimenter modelled it and asked her to imitate it.

Recognition phase

All six toys introduced in the presentation phase were placed in front of the child and the experimenter said,

(8) *Pokaż mi gdzie jest gryma.*
'Show me where the gryma is.'

waited for the child to respond, and provided feedback:

(9) *Tak, bardzo dobrze, to jest gryma.*
'Yes, very good, that's the gryma.'
(10) *Nie, to nie jest gryma, to jest …*
'No, that's not a gryma, that's …'

The same procedure was repeated with the remaining five items in the block.

Thus, the children heard each nonce noun at least nine times in a variety of constructions during the presentation and recognition phase, and each familiar word at least two times. Note that each nonce word was presented with an agreeing demonstrative (*cf.* (5) above) which provided a further clue about gender (in addition to the gender-typical offset).

Testing phase

The recognition phase was immediately followed by the testing phase, which consisted of three 'games', each designed to elicit the test nouns in a different case. The order of elicitation games was counterbalanced across sessions.

The genitive. The genitive case was elicited using the 'hiding game'. The experimenter took a toy out of a bag and said

(11) *Jest X!*
'Here's the X!'.

The child's task was to hide the toy in another bag as quickly as possible and say

(12) *Nie ma X-GEN*
'The X is gone!'[9]

If the child did not respond, he was prompted with the beginning of the test sentence (*Nie ma* ...); if he still did not respond, he was prompted again with the beginning of the sentence followed by the first syllable of the target word. The test began with three practice items (one for each gender), the first of which was modelled for the child. This was immediately followed by the six test words presented in random order.

The accusative. The accusative case was elicited using a procedure modelled on that described in Wittek & Tomasello (2005). The experimenter acted out a scene in which a toy first walked, and then was carried by a puppet, providing a running commentary on the events:

(13) *Zobacz! Idzie sobie lala. Idzie, idzie, idzie, zmęczyła się.*
'Look! The doll is walking. Walking, walking, walking, she got tired.'

---

[9] An anonymous JCL referee pointed out that we tested the accusative and dative in their most prototypical uses (to mark the patient and recipient respectively), and the genitive in a rather unprototypical use (after a negated verb). The use of the genitive with a negated verb is indeed unprototypical, and there is evidence that children have problems with it (Smoczyńska, 1985; Dąbrowska, 2001). However, the use of *nie ma* 'not have' plus the genitive to indicate absence is a special case which is probably unrelated (in children's grammars) to the use with other negated verbs; in any case, it is one of the earliest uses of the genitive.

The Kraków corpus (Smoczyńska, 1998) contains transcripts from five children under the age of 2. For four of the five children, the 'genitive of absence' appeared in the first 20 attested tokens of the genitive (at 17–18 months); the same was also true of two other uses, to mark possession and destination (with or without the preposition *do* 'to'). The genitive of absence accounts of about 11% of the children's uses of the genitive during this period. It is less frequent than the other early uses (possession, 27%; direction, 38%), but nevertheless was well established in all five children before age 2;0.

It is also worth noting that on a lexically specific level, the verb island *nie*+*ma*+GEN is considerably more frequent than the combinations used to elicit the other cases (*dać* 'give'+DAT, *nieść* 'carry'+ACC).

(14) [putting on a funny voice] *Ponieś mnie, kukiełko!*
'Carry me, puppet!'

(15) [makes puppet carry the doll] *A teraz kukiełka niesie lalę.*
'Now the puppet is carrying the doll.'

Then the experimenter made another toy 'walk' and asked the child to supply the last sentence.

(16) *Idzie sobie piesek. A teraz?*
'The dog is walking. And now?'

If the child did not respond, she was prompted with the beginning of the sentence:

(17) *Kukiełka niesie …*
'The puppet is carrying …'

If she still did not respond, she was prompted again with the beginning of the sentence followed by the first syllable of the target word. As in the previous game, there were three practice items, and the test items were presented in random order.

The dative. The dative case was elicited using the 'party game'. The experimenter explained to the child that all the toys had been invited to a party and each was to receive a present; the child's task was to decide which toy got which present. The first toys to arrive at the party were those used as the practice items. The experimenter said,

(18) *Najpierw na przyjęcie przyszły dziecko, miś i żyrafa.*
'The child, the teddy and the giraffe came to the party first.'

and placed the three toys in front of the child. Then she showed the child the first present and asked

(19) *Komu damy piłkę?*
'To whom should we give the ball?'

If the child did not respond, or responded by pointing, the experimenter prompted her with the beginning of the sentence

(20) *Piłkę damy …*
'We'll give the ball to …'

followed, if necessary, with the beginning of the sentence and the first syllable of the target word. The experimenter then placed the present in front of the toy indicated by the child and repeated the procedure with the next present. When all three toys used in the practice session had received their presents, they were put in the 'party room' and the experimenter proceeded to the actual test. The procedure used in the test was exactly the

same except that, in order to make the task easier for the child, the toys arrived two at a time. If at any time the child seemed to have forgotten what a particular toy was called, the experimenter modelled it one more time in a presentational construction (*To jest X* 'That's an X').

On completing the three 'games' the child helped the experimenter tidy up the toys and was asked if she wanted to continue playing. If the child was keen to continue, the next session was administered immediately afterwards; otherwise the experiment continued whenever she was available again (usually on the following day).

Coding

Each response was classified as one of the following:

1. Target: the expected ending for the given case-gender combination;
2. Zero: failure to inflect, i.e. use of the nominative rather than the required oblique form;
3. Overgeneralization: use of a gender-inappropriate ending for a given case (e.g. masculine ending with a neuter noun);
4. Other: failure to respond, use of an inappropriate case (e.g. genitive for dative), substitution of another noun.

Substitutions of the diminutive form of the same root for a simplex form, or vice versa, were counted as special cases of target, zero, or overgeneralization, as appropriate. Overgeneralizations were further subclassified according to which ending was overgeneralized (masculine, feminine, or neuter).

Problem cases. Because most endings serve as markers of more than one case (*cf*. Table 1), some responses are ambiguous:

1. The accusative neuter form is identical to the nominative, so it is impossible to distinguish correct accusative responses from zero responses.
2. Feminine nouns end in *-a* in the nominative; but *-a* also signals the genitive and accusative of some masculine nouns. Therefore, if a child produced a form ending in *-a* when presented with a feminine noun in a genitive or accusative context, the error could be regarded either as a failure to inflect or as an overgeneralization of the masculine ending.
3. Animate masculine nouns require *-a* in the accusative, but inanimate masculine nouns take *-Ø*. Thus, the bare stem in this case could be regarded either as a failure to inflect or as overgeneralization of the *-Ø* ending.
4. If a child added *-e* to a masculine noun in a dative context, the error could be regarded either as an overgeneralization of the feminine ending or as a case error (use of the locative ending instead of the dative).

In order to resolve such problem cases, the four categories mentioned above were applied in the order listed, that is to say

1. if a response could be regarded as correct, it was coded as target;
2. if a response could not be regarded as correct but could be considered a failure to inflect, it was coded as zero;
3. if a response was not correct but did involve adding an ending which could be regarded as an overgeneralization error, it was coded as an overgeneralization;
4. all other responses were coded as 'other'.

Thus, zero responses for the accusative neuter were coded as target (Principle 1); *-a* responses in the feminine genitive and accusative and bare-stem responses in the masculine accusative were coded as zero (Principle 2); and *-ę* responses in the masculine dative were coded as overgeneralizations of the feminine ending (Principle 3).

Another kind of problem case occurred when our participants used *-i/-y* in the dative condition. As indicated earlier, feminine nouns take *-i/-y* in the genitive and either *-i/-y* or *-ę* in the dative, depending on the phonological properties of the final consonant of the stem. All the feminine nouns used in the study required the *-ę* ending; therefore, the use of *-i/-y* could be regarded either as overgeneralizations of the other feminine ending or as case errors (use of the genitive instead of the dative). In accordance with principle 3, they should have been coded as overgeneralizations. There are, however, independent reasons for regarding them as case substitutions: the children who made such errors also tended to use the genitive ending with masculine nouns in dative contexts. Because of this, we coded such responses as 'other' errors.

Inevitably, this coding system introduces a certain degree of uncertainty, and therefore in our analysis of the results we concentrate on the target responses, which can be identified unambiguously.

RESULTS

This section is divided into five parts. We begin by presenting data on the overall number of different response types in each age group. In the following three subsections, we test the predictions of dual-mechanism and usage-based theories by examining the children's target responses and errors, and comparing their performance with that of adults. We conclude with a discussion of the role of frequency and phonological diversity at different stages of development.

General

All participants produced at least one correct genitive form and at least one correct non-neuter accusative form of a real word; and all but three children

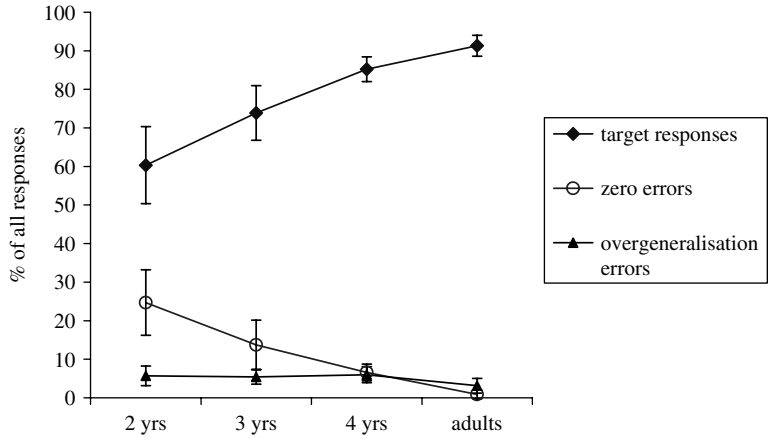

Fig. 1. Overall performance on real words.

(two two-year-olds and one three-year-old) produced at least one correct dative form of a familiar word. This shows that the children had understood the task and had the relevant syntactic knowledge: that is to say, they knew what case was required in the grammatical context in which the word was elicited. Furthermore, even in the youngest group, the majority of the participants were able to supply the target inflection with at least one of the four nonce words in all conditions except the dative neuter, showing that they were at least minimally productive with eight of the nine inflections.

Figures 1 and 2 show the overall performance on real words and nonce words in each age group. Each response type was analysed using a one-way ANOVA, separately for real words and for nonce words. For real words, we found a steady increase in the number of target responses ($F(3, 69) = 18{\cdot}141$, $p < 0{\cdot}001$) and a steady decrease in the number of zero responses ($F(3, 69) = 14{\cdot}168$, $p < 0{\cdot}001$). Overgeneralization errors were rare and occurred with similar frequency in all age groups ($F(3, 69) = 1{\cdot}565$, $p = 0{\cdot}206$). For nonce words, we found a steady increase in the number of target responses ($F(3, 69) = 8{\cdot}097$, $p < 0{\cdot}001$), a steady decrease in the number of zero responses ($F(3, 69) = 8{\cdot}061$, $p < 0{\cdot}001$), and a steady *increase* in the number of OG errors ($F(3, 69) = 8{\cdot}750$, $p < 0{\cdot}001$).

Target responses: the role of case, gender and lexicality

The children's accuracy data (percentage of target responses) were analysed using (3) × (3) repeated measures ANOVAs, with the main factors of gender (masculine, feminine, neuter) and case (genitive, dative, accusative). Two separate ANOVAs (for real and nonce words) were carried out. The effect of lexicality was analysed separately by comparing performance on real and

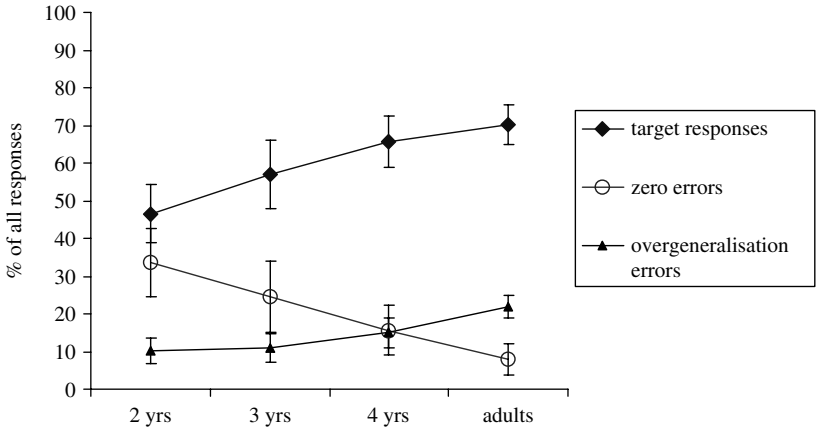

Fig. 2. Overall performance on nonce words.

nonce words for each case-gender combination. The F statistics reported below were computed using the multivariate Pillai's trace method. All significant main effects and interactions were followed up with t-tests. Eta squared (η^2) was used as a measure of effect size.

Since the data were collected from two-, three- and four-year-olds the factor of age could also have been included in the analysis. However, we decided to exclude it, for two reasons: (i) it is not theoretically interesting, since the competing theories do not make contrasting claims with respect to age; and (ii) floor and ceiling effects occurred in certain combinations of gender and case in the youngest and the oldest group, making parametric statistical analyses problematic. In order to illustrate age-related changes, we present the descriptive statistics separately for each age level in tables 6 and 7 below.

Real words

In Table 8, we present the mean, median, standard deviation and range in the number of target responses for all combinations of case and gender at each age level. Note that while even the youngest children performed relatively well on all inflections except the dative neuter, there were vast individual differences, particularly in the two-year-old group and in the scores for neuter words in all age groups.

Gender

The main effect of gender was significant and robust: $F(2, 55) = 77{\cdot}464$, $p < 0{\cdot}001$, $\eta^2 = 0{\cdot}74$. It resulted from neuter nouns (56% average accuracy)

TABLE 8. *Percent target responses* (*real words*)

		2-year olds ($N=18$)		3-year-olds ($N=20$)		4-year-olds ($N=19$)		All children ($N=57$)	
Case	Gender	Mean (*S.D.*)	Median [min–max]	Mean (*S.D.*)	Median [min–max]	Mean (*S.D.*)	Median [min–max]	Mean (*S.D.*)	Median [min–max]
Gen	Masc	79 (30)	100 [0–100]	88 (28)	100 [0–100]	99 (6)	100 [75–100]	89 (25)	100 [0–100]
	Fem	83 (31)	100 [0–100]	96 (17)	100 [25–100]	97 (8)	100 [75–100]	93 (21)	100 [0–100]
	Neut	40 (25)	50 [0–100]	58 (23)	50 [0–100]	65 (25)	75 [0–100]	54 (55)	50 [0–100]
Dat	Masc	58 (41)	75 [0–100]	68 (36)	75 [0–100]	83 (28)	100 [25–100]	70 (36)	75 [0–100]
	Fem	40 (37)	25 [0–100]	69 (38)	88 [0–100]	93 (11)	100 [75–100]	68 (37)	75 [0–100]
	Neut	6 (11)	0 [0–25]	19 (23)	13 [0–75]	41 (29)	50 [0–100]	22 (26)	25 [0–100]
Acc	Masc	63 (40)	75 [0–100]	86 (28)	100 [25–100]	96 (9)	100 [75–100]	82 (31)	100 [0–100]
	Fem	81 (36)	100 [0–100]	93 (16)	100 [50–100]	100 (0)	100 [100–100]	91 (23)	100 [0–100]
	Neut	93 (14)	100 [50–100]	90 (17)	100 [50–100]	93 (11)	100 [75–100]	92 (14)	100 [50–100]

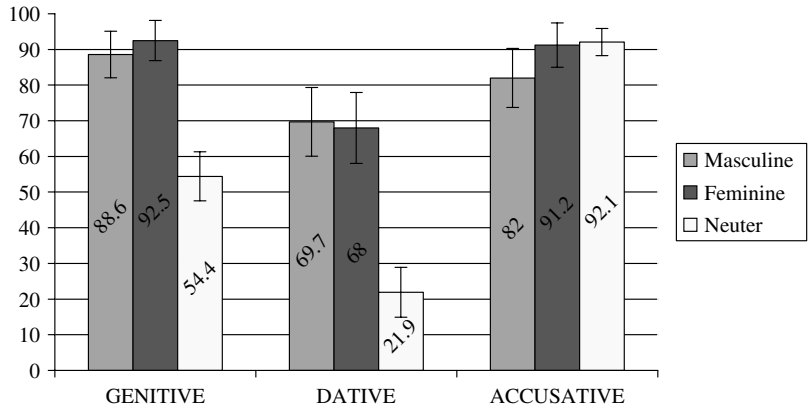

Fig. 3. Target responses on different levels of gender and case (real words). Note: Error bars show 95% confidence intervals for the mean.

being significantly more difficult than both masculine and feminine nouns (80% and 84% accuracy, respectively; see Figure 3). The difference between the latter two was not significant. The effect was qualified by a significant and robust interaction of gender by case: $F(4, 53) = 32{\cdot}338$, $p < 0{\cdot}001$, $\eta^2 = 0{\cdot}71$. Neuter nouns were significantly harder than masculine and feminine ones in the genitive and dative cases only; in the accusative case neuter nouns were significantly EASIER than masculines, and statistically no different from feminines. The good performance on the accusative neuter is attributable to the fact that this ending is identical to the nominative neuter ending, and hence the children could produce the target responses simply by repeating the form they had just heard. In addition, in the accusative case, masculine nouns were significantly more difficult than feminine nouns.

Case

The main effect of case was significant and robust: $F(2, 55) = 71{\cdot}095$, $p < 0{\cdot}001$, $\eta^2 = 0{\cdot}72$. It resulted from the dative case being significantly more difficult (53% average accuracy) than the genitive case (79%) which, in turn, was significantly more difficult than the accusative (89%). This pattern was qualified by the significant interaction of case and gender (see the previous subsection). While the dative was significantly harder than both the genitive and the accusative regardless of gender, the advantage for the accusative over the genitive appeared in neuter nouns only (see Table 8 and Figure 3).

Nonce words

The children's performance on nonce words is summarized in Table 9. Again, there are very large individual differences, especially in the younger children and in the neuter conditions.

Gender

The main effect of gender was significant: $F(2,55)=15{\cdot}863$, $p<0{\cdot}001$, $\eta^2=0{\cdot}37$. It resulted from neuter nouns (46% average accuracy) being significantly more difficult than feminines (58%) which, in turn, were harder than the masculines (67%). The effect was qualified by a significant interaction of gender and case: $F(4,53)=14{\cdot}409$, $p<0{\cdot}001$, $\eta^2=0{\cdot}52$. As with real words, neuter nouns were significantly more difficult than nouns belonging to other genders in the genitive and dative, but not in the accusative (where the neuter ending is identical to the nominative ending). The advantage for masculine over feminine nouns reached statistical significance in the dative case only (see Table 9 and Figure 4).

Case

The main effect of case was significant and robust: $F(2,55)=75{\cdot}982$, $p<0{\cdot}001$, $\eta^2=0{\cdot}73$. It resulted from the dative case being significantly more difficult (36% average accuracy) than the accusative case (63%) which, in turn, was significantly more difficult than the genitive (71%). This pattern was qualified by the significant interaction of case and gender (see the previous section). The dative was significantly harder than the genitive regardless of gender, but significantly harder than the accusative for feminine and neuter nouns only. In addition, the accusative was significantly harder than the genitive for masculine and feminine nouns, but not for neuters (see Table 9 and Figure 4).

Comparison of real words and nonce words

In Table 10, we compare the children's performance on real words and nonce words in each case-gender combination and give the results of 9 paired-samples *t*-tests assessing the significance of the difference between the two figures.

In all but one combination of case and gender there was a significant, moderate-to-large (following Cohen's convention: see Cohen, 1992) advantage for real words over nonce words. The differences were largest for feminine nouns and the accusative case. Only in the genitive neuter condition was there a small (statistically non-significant) advantage for nonce words over real words.

TABLE 9. *Percent target responses* (*nonce words*)

		2-year olds ($N=18$)		3-year-olds ($N=20$)		4-year-olds ($N=19$)		All children ($N=57$)	
Case	Gender	Mean (*S.D.*)	Median [min–max]	Mean (*S.D.*)	Median [min–max]	Mean (*S.D.*)	Median [min–max]	Mean (*S.D.*)	Median [min–max]
Gen	Masc	71 (27)	75 [25–100]	83 (32)	100 [0–100]	87 (19)	100 [25–100]	80 (27)	100 [0–100]
	Fem	69 (31)	75 [0–100]	80 (28)	88 [0–100]	82 (26)	100 [0–100]	74 (30)	75 [0–100]
	Neut	46 (37)	50 [0–100]	61 (38)	75 [0–100]	68 (27)	75 [0–100]	59 (35)	75 [0–100]
Dat	Masc	35 (34)	38 [0–100]	51 (40)	50 [0–100]	86 (19)	100 [50–100]	58 (38)	50 [0–100]
	Fem	25 (26)	25 [0–75]	35 (33)	25 [0–100]	61 (32)	75 [0–100]	40 (33)	50 [0–100]
	Neut	10 (21)	0 [0–75]	8 (18)	0 [0–75]	15 (21)	0 [0–75]	11 (20)	0 [0–75]
Acc	Masc	43 (36)	50 [0–100]	69 (33)	75 [0–100]	74 (29)	75 [0–100]	62 (35)	75 [0–100]
	Fem	56 (27)	50 [0–100]	56 (33)	63 [0–100]	62 (34)	75 [0–100]	58 (31)	75 [0–100]
	Neut	75 (23)	75 [25–100]	70 (35)	75 [0–100]	59 (40)	75 [0–100]	68 (34)	75 [0–100]

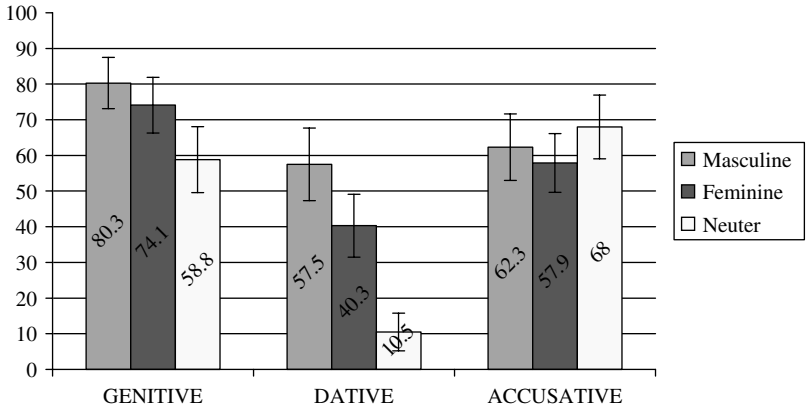

Fig. 4. Target responses on different levels of gender and case (nonce words). Note: Error bars show 95% confidence intervals for the mean.

Errors

Overgeneralization errors. Overgeneralization errors, as explained earlier, involve the use of a case ending which would have been appropriate for a noun belonging to a different class – for example, the use of the dative masculine ending with a feminine noun which occurs in a grammatical context which requires the dative. In our discussion of OG errors, we will consider two complementary questions, namely, which endings are overgeneralized, and which classes of words are targets of overgeneralization errors. Since the real words used in the experiment were not fully matched for frequency, it is difficult to interpret overgeneralization errors involving real words. In what follows, therefore, we will only consider the children's errors on nonce-word stimuli.

Table 11 shows the number of overgeneralizations of masculine, feminine and neuter endings in each case. As we can see, the children overgeneralized all endings, but masculine endings were overgeneralized more frequently than the others: 72% of all OG errors involved inappropriate use of a masculine ending. This figure may well be an underestimate: uses of the *-a* ending with feminine nouns in the genitive and accusative case, which we have classified as zero errors (since the nominative form also ends in *-a*) could be overextensions of the masculine ending. The number of overgeneralizations of the feminine dative ending is probably also somewhat higher: as pointed out earlier, some of the uses of *-i*/*-y* with feminine nouns that take *-e* could have been overgeneralizations of the other feminine ending rather than case errors.

TABLE 10. *Comparison of performance on real words and nonce words*

Case	Gender	Real words	Nonce words	*t*-test value	Sig.*	Effect size (Cohen's *d*)
Genitive	masc	89	80	3·029	0·004	0·32
	fem	93	74	6·497	<0·001	0·72
	neut	54	59	−1·299	0·199	−0·14
Dative	masc	70	57	3·221	0·002	0·33
	fem	68	40	7·193	<0·001	0·78
	neut	22	11	3·380	0·001	0·49
Accusative	masc	82	62	5·430	<0·001	0·59
	fem	91	58	7·584	<0·001	1·21
	neut	92	68	5·686	<0·001	0·93

* The reported significance values have not been corrected for multiple comparisons.

TABLE 11. *Children's overgeneralizations of masculine, feminine, and neuter endings to nonce words belonging to other genders*

Overgeneralized ending	Genitive	Dative	Accusative	All cases
Masculine	5	116	57	178
Feminine	12	18	16	46
Neuter	—	24	—	24
All endings	17	158	73	248

Note: The symbol '—' indicates that the relevant error could not occur because of the coding system used.

Table 12 gives information about the classes of words which are targets of overgeneralization errors. It is clear that while such errors occur with words of all three genders, they are considerably more frequent with neuter nouns than with masculine or feminine nouns.

Thus, the masculine declension has a positive 'trade balance' (masculine endings are often 'exported' to other genders, but feminine and neuter endings are rarely used with masculines); the neuter declension is a net 'importer'; and the feminine declension exports about as much as it imports.

Other errors. Most errors classified as 'other' involved substitution of a familiar word for the nonce word or failure to respond at all. In addition, the children sometimes used the genitive inflection instead of the dative, producing forms such as *lima* for *limowi* (masculine), *grymy* for *grymie* (feminine), and *toska* for *tosku* (neuter). As pointed out earlier, with feminine nouns such errors cannot be distinguished from overgeneralizations, since the same ending, *-i/-y*, is used to mark both cases with some nouns. However, the genitive and dative forms of masculine and neuter nouns are

TABLE 12. *Number of overgeneralization errors with nonce words of each gender (children)*

Gender of overgeneralized nonce word	Genitive	Dative	Accusative	All cases
Masculine	3	23	4	30
Feminine	1	50	0	51
Neuter	13	85	69	167
All words	17	158	73	248

TABLE 13. *Percent target responses obtained by adults and four-year-olds*

			4 year-olds ($N=19$)		Adults ($N=16$)	
Lexicality	Case	Gender	Mean (*S.D.*)	Median [min–max]	Mean (*S.D.*)	Median [min–max]
Real words	Genitive	Masc	99 (6)	100 [75–100]	100 (0)	100 [100–100]
		Fem	97 (8)	100 [75–100]	100 (0)	100 [100–100]
		Neut	65 (25)	75 [0–100]	86 (20)	100 [50–100]
	Dative	Masc	83 (28)	100 [25–100]	95 (10)	100 [75–100]
		Fem	93 (11)	100 [75–100]	100 (0)	100 [100–100]
		Neut	41 (29)	50 [0–100]	52 (37)	38 [0–100]
	Accusative	Masc	96 (9)	100 [75–100]	100 (0)	100 [100–100]
		Fem	100 (0)	100 [100–100]	100 (0)	100 [100–100]
		Neut	93 (11)	100 [75–100]	89 (16)	100 [50–100]
Nonce words	Genitive	Masc	87 (19)	100 [25–100]	97 (9)	100 [75–100]
		Fem	82 (26)	100 [0–100]	78 (24)	75 [25–100]
		Neut	68 (27)	75 [0–100]	66 (22)	63 [25–100]
	Dative	Masc	86 (19)	100 [50–100]	92 (20)	100 [25–100]
		Fem	61 (32)	75 [0–100]	75 (24)	75 [25–100]
		Neut	15 (21)	0 [0–75]	3 (9)	0 [0–25]
	Accusative	Masc	74 (29)	75 [0–100]	100 (0)	100 [100–100]
		Fem	62 (34)	75 [0–100]	81 (30)	100 [0–100]
		Neut	59 (40)	75 [0–100]	39 (27)	38 [0–75]

distinct, and it is clear that the children did not fully differentiate the two cases. Genitive-for-dative substitutions account for 10% of the children's responses in the dative condition (10% with masculine nouns, 14% with feminines, and 5% with neuters).

Adult controls

The results for adult participants are presented in Table 13. The table also includes data for the oldest children (i.e. the four-year-olds) for the purposes of comparison. Overall, adults supplied the target endings somewhat

more frequently than the four-year-olds, achieving higher accuracy in 7 out of the 9 inflections with real words and 5 out of 9 with nonce words. It is noteworthy, however, that the differences between the two groups were very small, especially for masculine and feminine real words, where the four-year-olds' performance was close to perfect. What is even more notable is that in some conditions it was the children who achieved higher scores. This occurred with feminine nonce words in the genitive and with neuter nonce words in all three cases, as well as with real neuter words in the accusative.

The possibilities of statistical analysis of the adult data are limited due to the ceiling effect. A series of Mann–Whitney tests was carried out to compare the performance of four-year-olds and adults in each of the 9 experimental conditions, separately for familiar words and nonce words. Seven differences approached or reached statistical significance ($p < 0{\cdot}10$). Adults were better than children on genitive neuter words ($p = 0{\cdot}010$), dative feminine words ($p = 0{\cdot}029$), genitive masculine nonce words ($p = 0{\cdot}053$), and accusative masculine ($p < 0{\cdot}001$) and feminine ($p = 0{\cdot}044$) nonce words. On the other hand, children were better than adults on dative neuter ($p = 0{\cdot}049$) and accusative neuter ($p = 0{\cdot}098$) nonce words. However, only the advantage of adults over children on accusative masculine nonce words remained significant once the Bonferroni correction for multiple comparisons was made.

Since it was noted that adults tended to perform worse than children in all three neuter nonce word conditions, another analysis, which grouped nonce words by gender (across the three cases) was carried out. This confirmed the trend. While adults were significantly better than children on masculine nonce words (96% vs. 82% average accuracy, $p < 0{\cdot}001$ on the Mann–Whitney test) and no different on feminine nonce words (78% vs. 68%, $p = 0{\cdot}180$), they were significantly worse on neuter nonce words (36% vs. 47%, $p = 0{\cdot}016$).

Adult performance was far from perfect, especially on feminine and neuter nonce nouns, and, unexpectedly, on real neuter nouns, where the number of target responses ranged from 52% in the dative condition to 89% in the accusative condition. The relatively poor performance on real neuter nouns was due largely to a tendency to substitute a related word for the test word. All the animate neuter nouns used in the experiment have morphologically related masculine forms derived by adding the affix *-ak*: for example, alongside *zwierz-ę* 'animal', *ciel-ę* 'calf', *piskl-ę* 'chick' we have the derived forms *zwierz-ak, ciel-ak, piskl-ak* with more or less the same meanings. Such substitutions account for 14% of all adult responses to real neuter nouns (8% in the genitive and accusative, 25% in the dative), and they were always inflected correctly. Thus, actual errors account for 11% of all responses to neuter nouns (6% in the genitive, 3% in the accusative, and 24% in the

TABLE 14. *Adults' overgeneralizations of masculine, feminine, and neuter endings to nonce words belonging to other genders*

Overgeneralized ending	Genitive	Dative	Accusative	All cases
Masculine	4	57	31	92
Feminine	10	16	8	34
Neuter	—	0	—	0
All endings	14	73	39	126

Note: The symbol '—' indicates that the relevant error could not occur because of the coding system used.

TABLE 15. *Number of overgeneralization errors with nonce words of each gender (adults)*

Gender of overgeneralized nonce word	Genitive	Dative	Accusative	All cases
Masculine	2	4	0	6
Feminine	0	15	0	15
Neuter	12	54	39	105
All words	14	73	39	126

dative); the majority of these are overgeneralizations of masculine endings. The fact that participants were most likely to resort to this strategy in the dative case, which was by far the most difficult for adults and children alike, strongly suggests that they were trying to avoid inflecting the neuter nouns. (Children also occasionally used morphologically complex masculine equivalents of the neuter stimuli; however, such substitutions were quite rare, accounting for less than 2% of their responses to neuter nouns. As in adults, the majority – 83% – of the substitutions occurred in the dative case.)

Tables 14 and 15 provide more detailed information about overgeneralization errors on nonce words made by the adult participants. As can be seen from these figures, the pattern of error is very similar to that seen in the children: the adults also predominantly overgeneralized masculine endings (73% of all errors), and were most likely to make OG errors with neuter nonce words (83% of errors). What is most striking about the adult data is the relatively high proportion of overgeneralizations (22% of all responses to nonce words, compared to 12% in the children; for real words, the relevant figures are 3% for adults and 6% for children).

TABLE 16. *Correlations between frequency and phonological diversity and performance on the nonce word inflection task* (*Spearman's rho,* N = 9)

	2-year-olds	3-year-olds	4-year-olds	adults
token frequency	0·817**	0·733*	0·383	0·400
type frequency	0·667*	0·683*	0·500	0·533
attested diversity	0·580	0·689*	0·740*	0·773*
potential diversity	0·068	0·305	0·865**	0·898***

*** Correlation is significant at the 0·001 level (2-tailed).
** Correlation is significant at the 0·01 level (2-tailed).
* Correlation is significant at the 0·05 level (2-tailed).

The relationship between productivity, frequency and phonological diversity across age groups

In this subsection, we examine in more detail the two factors that are most relevant for usage based theories – frequency and phonological diversity – with a view to determining how well they predict productivity at different stages of development. In earlier discussions we introduced two measures of frequency (type frequency, i.e. the number of different nouns that a particular inflection occurred with, and token frequency, i.e. its overall frequency) and two measures of phonological diversity (attested diversity, which was calculated on the basis of words which occurred with a particular inflection in a sample of parental utterances, and potential diversity, based on the number of nouns in the corpus which require the inflection, irrespective of whether they actually occurred with it or not).

Table 16 summarizes the relationship between these four measures and performance on the nonce word task. These results must be viewed with caution: since we elicited 9 endings, there are only 9 pairs of data points for each correlation. Nevertheless, several of the correlations are highly significant, and the table suggests some interesting patterns. Since the data are not normally distributed, we report the statistics for Spearman's rho. We also computed Pearson's r, and obtained a similar pattern of results, although the values of the correlation coefficients tended to be somewhat lower, and the differences between age groups smaller.

Both type and token frequency are good predictors of productivity, at least for the younger children (two- and three-year-olds). For four-year-olds and adults, the correlations are not significant, but given the size of the correlation coefficient, this is likely to be a result of the small number of data points rather than absence of a relationship. As observed earlier (*cf.* Table 6), the type and token frequencies of the nine endings are very strongly correlated (rho = 0·95, $p < 0·001$), which precludes drawing any firm conclusions about the relative contribution of these factors; but our data do suggest that

the relationships may change with age. Token frequency is very strongly associated with productivity in the youngest children, but the strength of the relationship decreases quite sharply with age. Type frequency is also more strongly associated with productivity in the younger participants; however, for this measure, the decrease is much smaller, so that for the four-year-olds and adults, type frequency appears to be a better predictor of performance on the nonce word task than token frequency.

Whatever the relationship between type and token frequency, the steady decrease in the correlation coefficients for both of these factors suggests that frequency plays a more important role in the earlier stages of development. In contrast to this, the role of phonological diversity appears to increase with age. This is most striking for potential diversity, where our data suggest no relationship in the two-year-old group and a very strong relationship (rho = 0·898) in adults.

DISCUSSION

Frequency

Given the frequency of the nine inflections which were the object of this study, usage-based theories make the following predictions:

1. Performance should be better on masculine and feminine nouns than on neuters;
2. In the dative case, performance on masculine nouns should be better than on feminine nouns;
3. In the accusative case, performance on feminine nouns should be better than on masculine nouns;
4. Performance should be better with genitive and accusative inflections than with dative inflections.

As shown in the Results section, predictions 1, 2, and 4 have been confirmed: performance on masculine and feminine inflections was indeed better than on neuter inflections (except in the accusative, where the neuter form is identical to the nominative); in the dative case, performance on masculine nouns was better than on feminine nouns; and across all genders, performance in the genitive and accusative cases was better than in the dative case. However, the predicted advantage for feminines over masculines in the accusative was not found in nonce words, although children did perform better on real feminine words. In fact, the children were slightly more accurate with masculine nonce nouns in all three cases – although in the genitive and accusative the difference was very small and not statistically significant.

There are several possible explanations for the relatively good performance on masculine nouns. Masculine endings may be more productive than

feminine ones because the masculine class is phonologically more diverse, and phonological diversity, as suggested in the introduction, may encourage schema extraction (see below for further discussion). It is also possible that Polish speakers rely on 'product-oriented' schemas (*cf.* Bybee, 1995). Feminine nouns end in *-a* in the nominative; since *-a* is also used to mark the genitive form of masculine and neuter nouns and the accusative form of animate masculine nouns, speakers might feel that the citation form is already genitive-like or accusative-like and does not need to be modified any further. The tendency to avoid inflecting words which already sound as if they had the correct inflection is well-documented: for example, English-speaking children are most likely to leave out the plural ending with nouns ending in [s] or [z], and the past tense ending when the verb stem ends in [t] or [d]; likewise, German children often omit the plural marker with 'pseudo-affixed' stems, i.e. those which end in [n] or [e], because the pseudo-affixes are identical with true plural markers (Bybee & Slobin, 1982; Köpcke, 1998; Taylor, 2002).[10] Finally, the particularly good performance on the dative masculine inflection, in comparison with the other dative inflections, may be partly attributable to its phonological salience (unlike all the other endings, it is disyllabic) and morphological distinctiveness (it is one of the few endings which has only a single function).

It should be stressed that these explanations are not mutually exclusive: the poorer performance on feminine nouns may be due to a combination of factors. It is also possible that different factors dominate at different stages of development: the particular salience of *-owi* in comparison with the other dative endings and reliance on product-oriented schemas are likely to have a greater impact on the performance of the youngest participants, while the performance of older children and adults may depend more on the phonological structure of the class.

Phonological diversity

Our discussion so far has focused on how productive each of the nine inflections was within its own domain of application. We found that participants supplied the target forms most reliably with nouns requiring masculine endings, although performance on feminines was almost as good; performance on neuter nouns, on the other hand, was considerably worse than on both masculine and feminine nouns.

[10] If this explanation is correct, then the advantage for masculines over feminines should disappear in the instrumental, where the citation form cannot be construed as a noun bearing the masculine ending.

Another aspect of productivity is how often a particular inflection is extended beyond its usual domain of application. In this respect, masculine inflections are considerably more productive than feminine and neuter inflections, accounting for over 70% of all overgeneralization errors with nonce words. Overgeneralizations of feminine endings accounted for about 20% of the children's errors and 27% of the adults', and neuter endings were rarely used with nouns of other genders.

Thus, masculine inflections are productive both in their own domain of application and outside of it; feminine inflections are productive in their own domain of application, and only occasionally overgeneralized; and neuter inflections are only weakly productive in their own domain of application, and almost never overgeneralized. As explained earlier, masculine and feminine inflections apply to larger classes of nouns and occur more frequently in the input than neuter inflections; and the domains of application of masculine inflections are more diverse than those of feminine inflections, which in turn are more diverse than neuter nouns. The pattern of results we observed suggests that an ending's frequency predicts how reliably children will apply it to nouns within its domain of applicability, i.e. to forms which resemble previously learned exemplars, while phonological diversity predicts how likely they are to generalize it to dissimilar nouns, including nouns belonging to other classes (and presumably also non-canonical nouns belonging to the same class).

Our results also indicate that the relative weight of frequency and diversity as determinants of productivity changes in the course of development. Frequency, especially token frequency, appears to be an excellent predictor of younger children's performance, but its role diminishes with age; on the other hand, the role of phonological diversity, especially potential diversity, increases with age. This suggests that learners' initial generalizations may be low-level schemas which apply to clusters of phonologically similar forms, while more general schemas which apply to more diverse classes are a feature of relatively mature linguistic systems.

This explanation makes sense of an otherwise puzzling finding: while performance generally improved with age, the number of overgeneralization errors actually increased (see below for further discussion). Additional support for such an account can be found in Dąbrowska's (2006) analysis of children's performance on simplex and diminutive nouns (which was based on the same data as the analyses in this paper). Dąbrowska found that young children are better at inflecting diminutives and diminutive-sounding nonce words (for which they are likely to have low-level schemas, since they are phonologically similar to real diminutive forms), in spite of the fact that they are more complex. However, the diminutive advantage decreased with age and was absent in adults, which suggests that mature speakers increasingly rely on more general rules.

Lexicality

We found a strong effect of lexicality: all age groups (including adults) performed better on real words than on nonce words. The average difference in the scores amounted to 17% and was very similar across all ages. There are at least two (mutually non-exclusive) explanations of this effect. Poorer performance on nonce words may be attributable to the additional processing demands required to use a newly learned word. Alternatively, the better performance on real words may be due to the fact that at least some inflected forms were retrieved from memory as ready-made units.

If the first explanation is correct, then we would expect participants to achieve higher scores on real words in all inflections, except when performance is at ceiling or at floor. This was clearly not the case: the advantage for real words was much greater for feminine nouns than for masculines and neuters, and there was no advantage for real words in the genitive neuter condition. Thus, it seems that the lexicality effect is at least partially attributable to the fact that the inflected forms of some real words are available as prefabricated units.[11]

Regularity

Our results suggest that speakers are most likely to generalize affixes which apply to large, phonologically diverse classes, and that they probably store many regularly inflected forms. Both of these findings are compatible with usage-based approaches and problematic for the dual mechanism theory. Furthermore, contrary to a fundamental tenet of this theory, regularity turned out to be a poor predictor of generalizability. Participants in all age groups readily applied the irregular masculine inflection to novel as well as familiar nouns. In contrast, performance was very poor on the genitive neuter and especially the dative neuter, in spite of the fact that these inflections are almost completely regular.

Syntactic errors

Smoczyńska (1985) points out that case-marking errors (i.e. use of a gender-inappropriate ending or the nominative form) are considerably more frequent than case substitutions (use of an oblique case form in a

[11] As explained in the Method section, real masculine and feminine nouns were matched for frequency, but the neuter words were less frequent. These differences in frequency could explain the difference in performance on feminine and neuter nouns, but not the differences between masculines and feminines, or the absence of lexicality effects in the genitive neuter.

grammatical context requiring another case), showing that children learn the basic syntactic distinctions of their language before they master the details of the inflectional system. She also notes one interesting exception to this tendency: some Polish-speaking children go through a stage during which they often use genitive-marked forms in grammatical contexts requiring the dative case. This was confirmed by our study. We found that syntactically inappropriate uses of oblique case forms were indeed very rare, except that many of the younger children sometimes used the genitive inflection when the dative case was required.

It seems that this failure to differentiate between the two cases is due to two factors. We have already noted a partial overlap in marking: some feminine nouns take *-i/-y* in both the dative and the genitive. There is also some semantic overlap, in that both cases can be used to mark possessive relationships, as in the following example:

(21) (a) *Robert zepsuł Piotrowi wiertarkę.*
Robert broke Piotr-DAT drill-ACC
(b) *Robert zepsuł Piotra wiertarkę.*
Robert broke Piotr-GEN drill-ACC
'Robert broke Piotr's drill.'

The two sentences are not fully synonymous: (21a) implies that Piotr was affected by the action in some way (for example, he is now unable to use the drill, or he is upset because it was a prized possession), while (21b) does not carry such implications. However the difference is quite subtle and presumably takes considerable time to master.

It is interesting to note that the opposite error (use of the dative in grammatical contexts requiring the genitive) did not occur in our data, although such substitutions are attested in the literature (see Smoczyńska, 1985). This could be due to the fact that dative relationships are semantically more complex than those coded with the genitive case; however, a comparison with the acquisition of case marking by German-speaking children suggests that semantic complexity is unlikely to be the main cause of Polish children's errors. The German case-marking system is very similar to Polish in the relevant respects: the dative case prototypically marks recipients, addressees, and experiencers, but can also indicate a possessive relationship; and the feminine dative and genitive are formally identical (for all feminine nouns, not just a subset, as in Polish). The critical difference appears to be frequency: in German, the genitive case is considerably less frequent than the dative. As a result, the German genitive is mastered late (Mills, 1985), and German-speaking two- and three-year-olds do not substitute genitive forms for the dative (Wittek & Tomasello, 2005). It would seem, therefore, that Polish children's tendency to use genitive inflections where the dative case is required is attributable to competition

from a semantically similar and much better entrenched form rather than to the semantic complexity of the dative.

Developmental trends and the adult system

We saw that productivity with case inflections develops early: the majority of the two-year-olds were already productive with all inflections apart from the dative neuter. Not surprisingly, performance improves with age, and by age 4, the children's scores are only slightly lower than those of the adult participants. However, the nonce word task revealed some exceptions to this general trend. Performance in the accusative neuter condition declined steadily with age, from 75% target at 2;6 to 39% in the adults. During the same period, the proportion of overgeneralization errors on nonce words increased from 10 to 22%. Finally, there was no improvement in performance in the dative neuter condition: all age groups had difficulty with this inflection, and the adult participants supplied the target ending in only 3% of the trials.

The first two developments are clearly related to each other. The increase in the number of overgeneralization errors is most plausibly interpreted as evidence of increasing productivity with masculine and feminine inflections. We saw earlier that younger children often leave unfamiliar nouns uninflected. This is unlikely to be due to lack of knowledge about which case is required in a given grammatical context, since they reliably inflect familiar nouns, especially in naturalistic settings (*cf.* Smoczyńska, 1985; Dąbrowska, 2001). It seems, then, that their relatively poor performance on nonce words is attributable to the fact that their schemas are less entrenched than adults', and possibly also less general (see the subsection on phonological diversity). Older participants, on the other hand, have more productive schemas, and tend to overgeneralize when they cannot access the correct inflection. As a result, OG errors displace zero errors in the course of development.

As explained in the introduction, for neuter nouns, the accusative neuter form is the same as the nominative. We can be reasonably certain, therefore, that the two-year-olds' good performance with this inflection is attributable not to their knowledge about the accusative neuter form, but to lack of knowledge: to produce the target response in this case, they merely had to repeat the form they had just heard. Older children gradually learn about the accusative neuter; but at the same time, they develop stronger and/or more general schemas for masculine and feminine inflections, and sometimes overgeneralize these to neuter nouns. Thus, performance on accusative neuter nonce words declines with age because (i) the neuter schema is relatively weak and (ii) there is growing competition from masculine and feminine schemas.

Perhaps the most surprising of our findings is the poor adult performance on neuter inflections, especially the dative neuter.[12] We argued earlier that this is attributable to the fact that they apply to a relatively small class of low phonological diversity. In the dative, the problem is compounded by two additional factors. First, the dative case is relatively infrequent, accounting for less than 2% of the noun tokens in the input (*cf.* Table 7). Secondly, the vast majority of neuter nouns are inanimate, while the dative case, because of its meaning, is used predominantly with animate nouns. As a result, learners hear relatively few exemplars of neuter nouns with the dative inflection, and do not become fully productive with it.

Still, the finding that adults have not mastered a very simple and almost completely regular part of the morphological system is bound to provoke some scepticism, so it is important to consider the possibility that the poor performance on neuter nonce words may be attributable to some property of the words themselves rather than the participants' knowledge, or lack of knowledge, about neuter inflections. As explained in the Method section, the neuter words were all phonotactically legal and ended in *-o*, the most typical ending for neuters, so it is unlikely that the effect is attributable to their phonological properties. However, unlike most neuter nouns, they referred to animals. Since animals are normally either male or female, it is possible that the participants in our study attributed a particular (natural) gender to the toys they were presented with and used the ending appropriate for the corresponding grammatical gender: the masculine ending if they thought the animal was male and the feminine ending if they thought it was female. Thus, the participants' behaviour could have been influenced by properties of the referent, and specifically, an unwillingness to assign neuter gender to a noun designating an animal.

There are two points to bear in mind in this connection. First, although natural gender is a reasonably good predictor of grammatical gender, it is not entirely reliable, and when the two are in conflict, it is the latter that determines the choice of ending. There were two linguistic cues to gender in our study: the phonological form of the nominative, and an agreeing demonstrative. The former, like natural gender, is a probabilistic cue; but the latter is fully reliable. Thus, to the extent that the participants were relying on natural gender rather than grammatical gender, they were violating the rules of their language. Secondly, there is converging evidence from a larger nonce-word study which confirms that Polish adults

[12] The problem with neuter inflections is least apparent in the genitive case, where the adult participants supplied the target ending in 66% of the opportunities. This figure, however, is likely to be an overestimate of their productive potential: because the genitive neuter and the genitive masculine ending for animate nouns are the same (*cf.* Table 1), some of the responses that were coded as target could in fact have been overgeneralizations of the masculine ending.

are only weakly productive with neuter endings (Dąbrowska, 2004). The two experiments described by Dąbrowska used different nonce words and a different elicitation method, and, critically, the neuter words had inanimate referents; yet her findings were very similar: participants had difficulty inflecting neuter nouns, but nearly always supplied the target form with masculine and feminine nouns. Thus, while confusion about gender might have contributed to our participants' difficulties with neuter nouns, it cannot be the main factor responsible for their poor performance.

CONCLUSION

The results reported above suggest that both frequency and phonological diversity affect productivity, albeit is different ways. Endings which occur frequently become better entrenched, and, as a result, are reliably applied within their domain of applicability, that is to say, to forms which resemble previously learned exemplars. Endings which apply to phonologically more diverse classes, on the other hand, are more likely to be generalized to forms which are dissimilar to those which occurred in the learner's experience, including forms which do not belong to the inflection's usual domain of application. We also found that frequency, especially token frequency, is the best predictor of younger children's productivity, while phonological diversity appears to be more relevant in later stages of development. These findings suggest that learners' initial generalizations are phonologically specific schemas, and that more general rules emerge later in development, possibly as a result of generalization over the early low-level schemas rather than actual exemplars. We also found a strong effect of lexicality, in both children and adults, which suggests that many inflected forms are available to speakers as ready-made units, regardless of whether they can also be produced by applying a rule. All of these findings are consistent with usage-based models. On the other hand, we found no support for the dual mechanism theory: not only was there no sharp dissociation between regulars and irregulars, but regularity turned out to be a very poor predictor of productivity.

REFERENCES

Bańko, M. (2000). *Inny słownik języka polskiego*. Warszawa: Wydawnictwo Naukowe PWN.

Bodnarowska, J. (1962). Problematyka doboru końcówek *-a/-u* w dopełniaczu l.p. rzeczowników męskich. *Język Polski* **42**, 29–49.

Bybee, J. L. & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language* **58**, 265–89.

Bybee, J. L. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes* **10**, 425–55.

Clahsen, H. (1999). Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behavioral and Brain Sciences* **22**, 991–1060.

Clahsen, H., Aveledo, F. & Roca, I. (2002). The development of regular and irregular verb inflection in Spanish child language. *Journal of Child Language* **29**, 591–622.
Cohen, J. (1992). A power primer. *Psychological Bulletin* **112**, 155–9.
Dąbrowska, E. (2001). Learning a morphological system without a default: the Polish genitive. *Journal of Child Language* **28**, 545–74.
Dąbrowska, E. (2004). Rules or schemas? Evidence from Polish. *Language and Cognitive Processes* **19**, 225–71.
Dąbrowska, E. (2005). Productivity and beyond: mastering the Polish genitive inflection. *Journal of Child Language* **32**, 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**, 120–35.
Elman, J. L., Bates, E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996). *Rethinking innateness: a connectionist perspective on development*. Cambridge, MA: MIT Press.
Köpcke, K.-M. (1998). The acquisition of plural marking in English and German revisited: schemata versus rules. *Journal of Child Language* **25**, 293–319.
Kottum, S. E. (1981). The genitive singular form of masculine nouns in Polish. *Scando-Slavica* **27**, 179–86.
Langacker, R. W. (1991). *Concept, image and symbol: the cognitive basis of grammar*. Berlin: Mouton de Gruyter.
Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (eds), *Usage-based models of language*. Stanford, CA: CSLI Publications.
Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R. & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology* **29**, 189–256.
Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J. & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development* **57** (Serial number 228), 1–180.
Mills, A. (1985). The acquisition of German. In D. I. Slobin (ed.), *The crosslinguistic study of language acquisition. Volume 1: The data*. Hillsdale, NJ: Erlbaum.
Pinker, S. (1998). Words and rules. *Lingua* **106**, 219–42.
Pinker, S. (1999). *Words and rules. the ingredients of language*. London: Weidenfeld and Nicolson.
Plunkett, K. & Marchman, V. (1993). From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition* **48**, 21–69.
Rumelhart, D. E. & McClelland, J. L. (1986). On learning the past tenses of English verbs. In D. E. Rumelhart & J. L. McClelland (eds), *Parallel distributed processing: explorations in the microstructure of cognition. Volume 2: Psychological and biological models*. Cambridge, MA: MIT Press.
Slobin, D. I. (1985). Crosslinguistic evidence for the language-making capacity. In D. I. Slobin (ed.), *The crosslinguistic study of language acquisition. Volume 2: Theoretical Issues*. Hillsdale, NJ: Erlbaum.
Smoczyńska, M. (1985). The acquisition of Polish. In D. I. Slobin (ed.), *The crosslinguistic study of language acquisition. Volume 1: The data*. Hillsdale, NJ: Erlbaum.
Smoczyńska, M. (1998). Krakowskie dane językowe dzieci w systemie CHILDES. In M. Smoczyńska (ed.), *Studia z psychologii rozwojowej i psycholingwistyki*. Kraków: Universitas.
Taylor, J. R. (2002). *Cognitive grammar*. Oxford: Oxford University Press.
Tokarski, J. (2001). *Fleksja polska*. Warszawa: PWN.
Westfal, S. (1956). *A study in Polish morphology. The genitive singular masculine*. The Hague: Mouton.
Wittek, A. & Tomasello, M. (2005). German-speaking children's productivity with syntactic constructions and case morphology: local cues act locally. *First Language* **25**, 103–25.

APPENDIX

WORDS USED IN THE EXPERIMENT

(Note: The second form given is the diminutive.)

Gender	Real words	Nonce words
Masculine	*baran/baranek* 'ram'	*pur/purek*
	zając/zajączek 'hare'	*lim/limek*
	ptak/ptaszek 'bird'	*czumas/czumasek*
	robak/robaczek 'bug'	*grut/grutek*
Feminine	*żaba/żabka* 'frog'	*zora/zorka*
	krowa/krówka 'cow'	*gryma/grymka*
	małpa/małpka 'monkey'	*ksiuda/ksiudka*
	ryba/rybka 'fish'	*klota/klotka*
Neuter	*zwierzę/zwierzątko* 'animal'	*toso/tosko*
	cielę/cielątko 'calf'	*żulo/żulko*
	kurczę/kurczątko 'chicken'	*klimo/klimko*
	pisklę/pisklątko 'chick'	*prato/pratko*