EWA HAMAN
University of Warsaw

EARLY PRODUCTIVITY IN DERIVATION.
A CASE STUDY OF DIMINUTIVES IN THE ACQUISITION OF POLISH[*]

Early speech diary data of one Polish girl - Inka (1;0 – 2;2) were analyzed to establish the course of emergence of the first derivational category in the child speech. Diminutives which are in general very productive category in Polish and are especially frequent in child-directed speech (CDS) occurred to be formed productively by the child as early as at the age of 1;5. Five indicators of category productivity in the child's speech were used: (1) total number of diminutives; (2) spontaneous use of diminutives; (3) number of different formants used in DIM words; (4) number of different stems used by Inka in DIMs; (5) proportion of diminutives for which simple forms were found. First four of them showed systematic developmental changes indicating that about age 1;5 Inka started to form diminutives productively. Frequency of diminutives in Inka's speech was compared to that in adult's utterances directed to her. The comparison revealed that the profile of proportion of DIMs in Inka's speech does not reflect the quantitative characteristics of CDS, since Inka used proportionally more diminutives than adults from the age of 1;5.

Introduction

First words and first phrase constructions have always been carefully studied (e.g., Braine, 1963; Brown, 1973; Chomsky, 1965; Tomasello, 1992; Bloom, 2000), but for several reasons first complex words have attracted much less attention (e.g., Clark, 1993, 1995). The problem of acquisition of word meaning has mainly been viewed as establishing the relation between linguistic sign and surrounding reality (e.g., Bloom, 2000), whereas the acquisition of complex word meaning can be seen as discovering the systematic relation between linguistic signs. The meaning of a complex word is in fact defined by its relation to a base/simple word.

This relation is built into the structure of a complex word and could be interpreted as analogous (though not identical) to the phrase structure, which has always been so interesting for researchers. But while all languages make similar use of phrase struc-

---

[*] Address for correspondence: Ewa Haman, Faculty of Psychology, University of Warsaw, Stawki 5/7, 00-183 Warszawa, Poland. E-mail: meh@sci.psych.uw.edu.pl

tures (despite wide differences in syntactic rules in particular languages), word-formation (the use of word structure) seems to be rather an optional device. Some languages have very productive word formation (e.g., Polish or other Slavic languages), while others use it much more marginally (e.g., English). The long domination of English in developmental psycholinguistics is also responsible for the low interest in word formation (Derwing & Baker, 1986; Aitchison, 1994). There are not many studies on derivational development in languages with prolific word-formation (but note e.g., Czukowski, 1968 for Russian; Chmura-Klekotowa, 1971; Haman, 2002 for Polish).

Inflectional morphology treated as a reflection of syntactic rules has won much more interest than derivational morphology used for formation of new words. The typical example of the difference in the approach to these two kinds of analysis of word structure are the structures studied in the classical experiment of Berko (1958). Everybody knows the 'wug' test as a test of the creative use of plurals, but almost no one bears in mind that the experiment included testing of some derivational affixes.

Among the various word-formation rules, formation of diminutives has never had a special status. What exceptionally interesting could there be in these "very small" and "really nice" words which in many languages characterize child-directed speech, but are of no importance to language acquisition?

However, some new data point out the function that diminutives can serve in early language. Dressler and Merlini-Barbaresi (1999) claim this is very first device used by children in derivational morphology first in a fictive function. Kempe, Brooks and Pirott (2001) show that in some languages extensive use of diminutives in child directed speech can facilitate the acquisition of grammatical gender (which was shown for Russian and Spanish). Savickiene (1998) argues that Lithuanian children need diminutives to simplify the inflectional system (use of diminutives narrows the number of declinations).

But to make some use of diminutives children firstly have to be able to analyze their structure, to recognize the underlying derivational rule and then to apply it productively. How it starts? The analysis of very first diminutives in early speech of a Polish girl (age 1;0-2;2) presented in this paper undertakes to answer this question.

Formation of diminutives in Polish

Diminutives are a very productive derivational category in Polish (Grzegorczykowa, Puzynina, 1998a,b; Nagórko, 1998), especially in child-directed speech. Semantic and pragmatic functions of DIM in adult speech are interchangeable and often formally 'core' semantic diminutives are used as hypocoristics. The main grammatical category which can be diminutivized are nouns, but adjectives and adverbs are also diminutivized sometimes (*malutki – mały* DIM – small DIM; *szybciuśko – szybko* DIM – quickly DIM). Formally, it is also possible to coin DIM verbs but they are very rare (e.g. *płakusiać* derived from *płakać* – to cry; *brykuńciać* – from *brykać* – to frisk).

It seems that in child language nouns form the main and most interesting DIM group.

Gender of noun does not change with diminutivization and corresponding suffixes are used in masculine (M), feminine (F) and neuter (N) genders.

Formants (suffixes) – in 'core' diminutives
most frequent: **-k-** (singular; **-ki** plural for all genders)
-ek/i(y)k (M) → *piesek* – a dog DIM; *kocyk* – a blanket DIM; - *krzewik* – a shrub DIM
-ka (F) → *lampka* – a lamp DIM
-ko (N) → *gniazdko* – a nest DIM
others: **-uszek**, **-iszek**, **-aszek**, **-i(y)czek** (all M, but F and N forms also exist) –
    dzbanuszek – a jug DIM; *braciszek* – a brother DIM; *wujaszek* – an uncle
    DIM; *czajniczek* – a teapot DIM;

Formants (suffixes) in hypocoristics
-uś (M); **-usia** (F); **-usio** (M, N) → *Jacek* – *Jacuś (*proper name*)*; *babcia* –*babusia*
    (grandmather); *wnuk* – *wnusio* (grandson)
-cio; -cia; → *wuj* – *wujcio* (an uncle); *córka* – *córcia* (a daughter)
-unio; -unia; → *mąż* – *mężunio* (a husband); *mama* – *mamunia* (a mother)
-ulo; -ula; → *tata* - *tatulo* (a father); *mama* – *mamula* (a mother)
-uchna; -uchno → *żaba* – *żabuchna* (a frog)
-ątko → *dziewczę* – *dziewczątko* (a girl)
-o → *dziadek* – *dziadzio* (a grandfather)

Combining of formants
different DIM formants can be attached to the same root slightly changing the
meaning (usually adding some expressional meaning) → *piesek – piesuś – piesunio*
(a dog DIM); *żabka – żabcia – żabuchna – żabula – żabusia* (a frog DIM);
several formants can be attached to one root at a time (II level diminutivization):
mama (simplex – a mother) → *mamusia – mamusinka, mamusieńka,*
mamusienieczka.

The data

The analysis of speech diary material enabled finding the very first uses of di-
minutives. It would be extremly difficult to conduct an experimental study on the
productive use of diminutives with subjects of this age. Using diary data also made
possible contrasting the characteristics of child speech with child-directed speech
(henceforth: CDS).

The analyzed diary comes from the collection of Szuman diaries. These are natu-
ralistic speech data gathered in the 1960s in Kraków, Poland, by mothers (psychology
students) of children, over an exceptionally long period of time: the longest diary was
kept for about 7 years (Smoczyńska, 1998). Although the collection of speech diaries
in Polish[1] is extensive, there is only one starting before age 1. This is also the diary
which was kept very systematically, so both child and adult data are very reliable. This
is the diary of a girl Inka.

---

[1] All diaries edited by Smoczyńska are available on the CHILDES web page: www.poppy.psy.cmu.edu/
childes.

Table 1. File size - number of different word types used by Inka, number of tokens and type/token ratio for each age interval

age	1;00	1;01	1;02	1;03	1;04	1;05	1;06	1;07	1;08	1;09	1;10	1;11	2;00	2;01	2;02	total
types	66	55	59	45	99	168	182	218	324	534	746	619	353	432	472	2157
tokens	249	232	270	130	399	806	868	1022	1306	1885	2291	1902	860	1128	1389	14737
tp/tk ratio	0,27	0,24	0,22	0,35	0,25	0,21	0,21	0,21	0,25	0,28	0,33	0,33	0,41	0,38	0,34	0,146

Figure 1. Mean Length of Utterence in Inka's speech
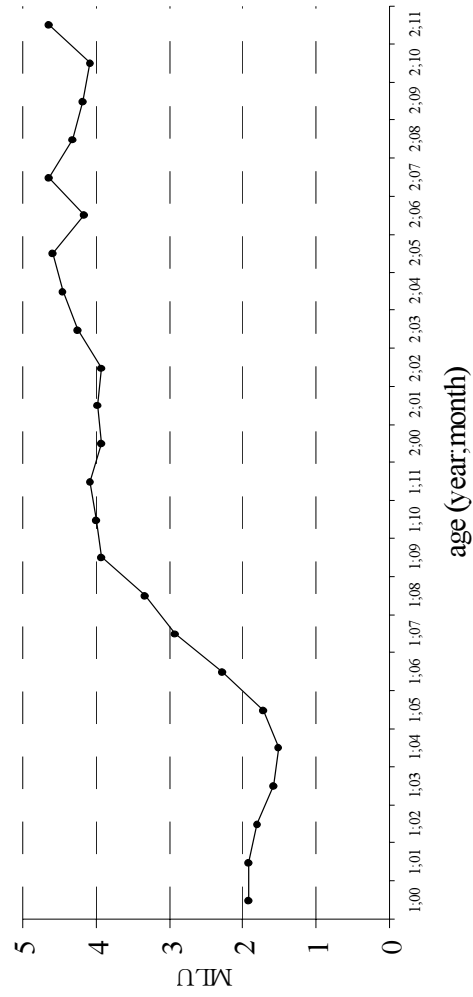
The diary was kept for over six years, starting at the age of 10 months (before the period of Inka's active speech) and ending at the age of 7 years and 8 months. Data were collected in one-month files (total number of files – 82) and no single file is omitted! It is also the richest diary – comprising 305.178 in all word forms, in which 56% are Inka's (170.574 forms). The period of 15 months – from 1; 0 to 2; 02 was chosen for the investigation of early diminutives in Inka's speech. Two earlier files (age 10 and 11 months) were excluded from the analysis, because they contained no diminutives. The chosen period includes first occurrences, first productive uses, massive "overuse" and the beginning of balanced usage of diminutives. The data were analyzed at one-month intervals.

This was also a period of crucial changes in other aspects of Inka's speech development. Figure 1. shows changes in MLU during the period 1;0-2;11. MLU was calculated by the MLU program, one of CLAN programs (McWhinney, 1995)[2]. The essential growth is seen from 1;06 to 1;10 (MLU rises from 2,28 to 3,93). After the age 2;02 the MLU rises slightly again, but previous analysis of diminutives in Inka's later speech (from 2;00 to 6;11) revealed no important developments in the category of diminutives at this age range (Haman, 2000), thus the present analysis was narrowed to the period lasting to the age 2;02.

For naturalistic data gathered in a home context it is very difficult to establish a constant amount of speech samples to be collected, especially in a phase of substantial developmental changes in child's speech (about the time of vocabulary spurt). Thus the monthly files differ significantly in size. Table 1. shows file size measured by number of type and token frequencies of all words used by Inka in each age interval, last column shows total number of different word types used and total number of words (tokens) calculated by FREQ. The sum of types from all intervals do not equal total number of types (in the last column) because the latter number was counted on cumulated data from the overall period. In the subsequent statistical comparisons a proportion of a category to file size was often used to avoid differences resulting from file size variation.

Data analysis

The analysis comprised all data files from the period 0;10-2;02. All diminutives used by Inka were extracted from frequency lists (made by FREQ). The two earliest files (ages 10 and 11 months) were excluded from the analysis, because there were no diminutives at all. Thus all subsequent analyses will concern the period 1;0-2;02. It occurred also that almost all diminutives found in Inka's speech were nouns. Only a few diminutivized words used by Inka were of a different word-class category. In the analyzed data only one adjective DIM was found (*malutki* – small DIM – in 5 different morphological forms) and two adverbs (*cieplutko* – warmly DIM; *prędziutko* – quickly DIM). No adjectives and adverbs were found before age 1;09 and no

---

[2] all programs for speech diaries analysis used for current analysis are available in CHILDES system – on the web at: www.poppy.psy.cmu.edu/childes/clan

Table 2. Number of DIMs (in types and tokens); repetitions and spontaneous uses of DIMs; numbers of different formants and stems used in DIMs for each age interval

age	1;0	1;1	1;2	1;3	1;4	1;5	1;6	1;7	1;8	1;9	1;10	1;11	2;0	2;1	2;2
DIM (types)	5	2	6	6	11	30	39	52	77	136	190	170	88	100	93
DIM (tokens)	8	5	14	14	35	98	176	240	313	514	681	524	223	257	43
DIM-repeated	5	0	5	2	6	11	12	2	14	16	14	13	18	10	11
DIM-spontanous	0	2	1	4	5	19	27	50	63	122	176	157	70	90	82
formants	4	2	4	3	4	10	10	12	11	21	22	22	17	21	14
stems	3	2	6	6	11	27	30	42	56	81	83	90	54	66	57

diminutivized verbs were found at all. Thus in fact all the following analyses and conclusions concern the category of nouns.

The main goal of the analysis was to establish the starting point of productive use of diminutives.

In the lexical derivational categories one should firstly look for innovations. The problem is that in the case of Polish diminutives the category is so productive itself that almost any potential innovation can be found as a correct, acceptable word in standard adult Polish. Therefore new DIM words in the child's speech almost never can be classified as innovations. This was the reason for adopting other indicators of derivational productivity of diminutives in Inka's speech. Five indicators of productivity were used and are described in detail in the following sections.

All calculations were made in both type and token frequencies, although type frequency seems to be a much more adequate statistic for productivity assessment (Clark, 1993). Type frequency tells us how many different words of the category were used, while token frequency informs only about the number of uses. Thus the same word repeated many times results in high token frequency of the category, but does not indicate its productivity, whereas many different words used only once each (resulting in high type frequency) may suggest that the category is productive in the child's language. Thus significance tests were computed on type frequency data.

The indicators of productivity are:

(a) proportion of DIMs in a child's speech

If words from any derivational category are acquired as unanalyzed units, analogously to simple words, there is no reason to expect any developmental changes in the frequency of this category except those resulting from overall vocabulary growth. Thus the proportion of frequency of an unproductive lexical category to all words used should be constant across age. In contrast, the increase in the category proportion can be seen as a sign of its development. A developing category should be almost from a certain point productive. The question is how to set the starting point of its productivity? The total number of DIMs and its proportion to file size do not give a direct answer to this question, but comparison with other indicators should help in solving the problem.

(b) spontaneous use of DIMs

In the early stages of lexical development it is very probable that many words used by a child are just repetitions of what the adult has just said. Such repetitions can hardly be treated as productive uses. Thus it is reasonable to differentiate repetitions and spontaneous uses. The increase of spontaneous uses of DIMs can be another indicator of category development and productivity.

 (c) number of different formants used in DIMs

DIM formation is realized by adding suffixes (formants) to a simple word stem. Giving the multiplicity of DIM formants in Polish, one can assume that an increase of the number of formants across age can be the next indicator of category productivity.

Table 3. Examples of Inka's diminutives. Simple words in square brackets were not found in the data (although Inka used diminutives derived from these words)

age	simplex	English translation	Diminutives (different types)	standard/ innovative DIM
1;6	dziadek	grandfather	dziaduś	standard
			dziadzia	standard
			dziadziuś	standard
	[ziemniak]	a potato	ziemniaczek	standard
1;7	dom	a house	domik	innovation
	[żaba]	a frog	żabka	standard
			żabunia	standard
1;8	kawa	coffe	kawka	standard
			kawusia	standard
	[chleb]	bread	chlebek	standard
			chlebuś	standard
1;9	Inka	girl's name	Inunia	standard
			Inuś	standard
			Inusia	standard
	auto	a car	autusio	standard
			autko	standard
1;11	kompot	compote	kompocik	standard
	lala	a doll	laluńka	standard
			lalunia	standard
			lalusia	standard
			laluś	standard
2;1	[stryj]	an uncle	stryjcio	standard
	[jabłko]	an apple	jabłuszko	standard
	ślimak	a snail	ślimaczek	standard
2;2	truskawka	a strawberry	truskaweczka	standard
	kapusta	a cabbage	kapustka	standard

Frequent use of only one or several formants could imply productivity of this particular formant but not category productivity. The use of a wide range of formants in the DIM category indicates the productivity of the category itself.

(d) number of different stems used in DIMs

The formation of derived words can be restricted to a small number of stems involved in the process, which could mean that the derivation is limited to particular lexical items and does not concern the same range of words as in an adult-like category. The more stems a child uses for DIM formation the more probable the process relates to the set of words analogous to that which can be diminutivized in adult language and thus can be seen as categorical, e.i. not restricted to particular words. Thus

Figure 2. Proportion of diminutives to total number of words used by Inka (types and tokens)
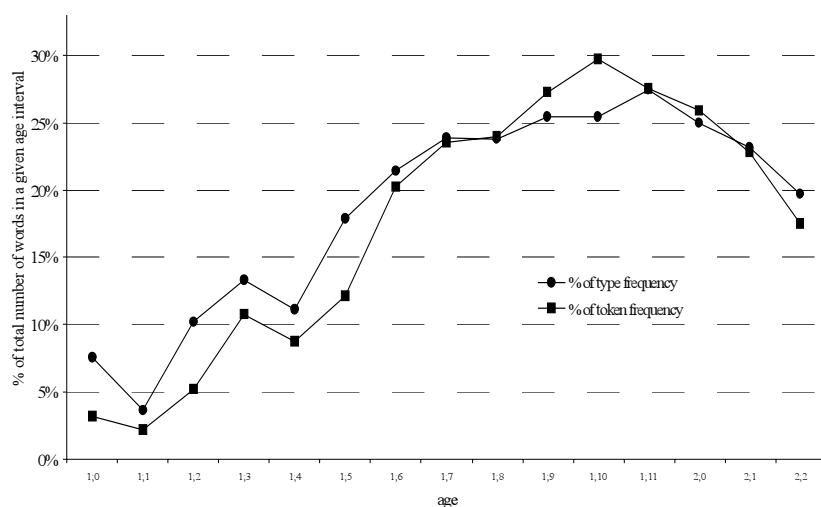
an increase of stems used in DIMs by a child means that he or she enlarges the scope of words used for diminutivization, which can indicate the development of the category and its productivity.

speech was found
een a derivational base
vative are present in a
on is much easier for a
's speech does not nec-
hild can still know the

a's speech will be de-
ll indicators (in 15 age
diminutives (or to file
f DIM used by Inka at

The proportion of DIMs (both in types and tokens) in Inka's speech increase with age (Figure 2) for ten consecutive age intervals and then declines slightly. Pearson correlation with age calculated for proportion of DIM type frequency in a file size on the data from all age intervals is significant (r=.83, N=15, p<.001). To check whether there are any crucial points in proportion changes a chi^2 test comparing DIMs type frequency with type frequency of all other words used by Inka was run.

The data were organized into three age intervals (frequencies were summed in the five-month intervals: 1;0-1;04; 1;05-1;09 and 1;10-2;02). There were two reasons for

Table 4. Summed frequencies of DIMs and all other word types in Inka's speech in three age intervals

	1;0-1;4	1;5-1;9	1;10-2;2
DIM type frequency	30	334	641
File size – DIM type frequency	294	1092	1981

employing this age division. Firstly, the three intervals are of the same length and secondly (and more importantly), it seems that the level of 20% (Figure 2) crossed between the age 1;5 and 1;6 may be critical for DIMs development (it is the level of DIMs proportion in CDS – see section 4.7), which will be discussed in detail in the following sections. The result of the chi^2 test is statistically significant – chi^2=37.8, df=2, p<.000001. Table 4. presents summed frequencies of DIM types and all other words types used in chi^2 tests.

Thus the overall proportion of DIMs in the total number of words changes with age. The direction of change – systematic increase to the age 1;10 – suggests that it is not only new word learning from the adult input but probably some rule facilitating the production of DIMs that is responsible for the change. Two additional chi^2 tests for 2x2 tables – comparing frequencies in pairs of two consecutive intervals revealed that the effect is due to the difference between intervals 1;0-1;4 and 1;5-1;9 (chi2=32.1, df=1, p<.00001). The difference between the second (1;5-1;9) and third (1;10-2;2) intervals was not significant (chi2=.53, df=1, p=.42). This suggests that crucial development of the category occurs about the age 1;5.

The effect is illustrated by the proportion of DIMs in the total number of words (Figure 2). If DIMs had been acquired as simple (unanalyzed) words from the adult input, then the proportion should had been constant. The alternative explanation of the observed changes could be analogous frequency changes of DIM words in Child-directed Speech – this possibility will be tested after presentation of all five productivity indicators.

Then increasing proportion of DIMs in the total number of words used across age (if not reflecting CDS) indicates productive use of DIMs in Inka's speech, from some point of the development. The results of chi^2 tests imply that the point is located about the age 1;5.

Spontaneous use of diminutives

All DIMs found in the frequency lists were checked in the context they were used. They were then classified as repetitions or spontaneous uses. If an adult (or other speaker) introduced the word (DIM) in the episode and then Inka repeated it (either in the same form or with inflectional case change), it was classified as repetition. If Inka herself introduced the word (DIM) in the episode it was counted as spontaneous use. In summing up spontaneous uses for a given age interval it seemed reasonable to count a word as spontaneous if it was used as such even once, since when the child is able to introduce the word in one episode then repetitions in other episodes might be interpreted as resulting from pragmatic factors in conversation rather then inability of spontaneous use.

Figure 3. Proportion of direct repetitions and spontaneous uses of diminutives to total number of diminutives in Inka's speech
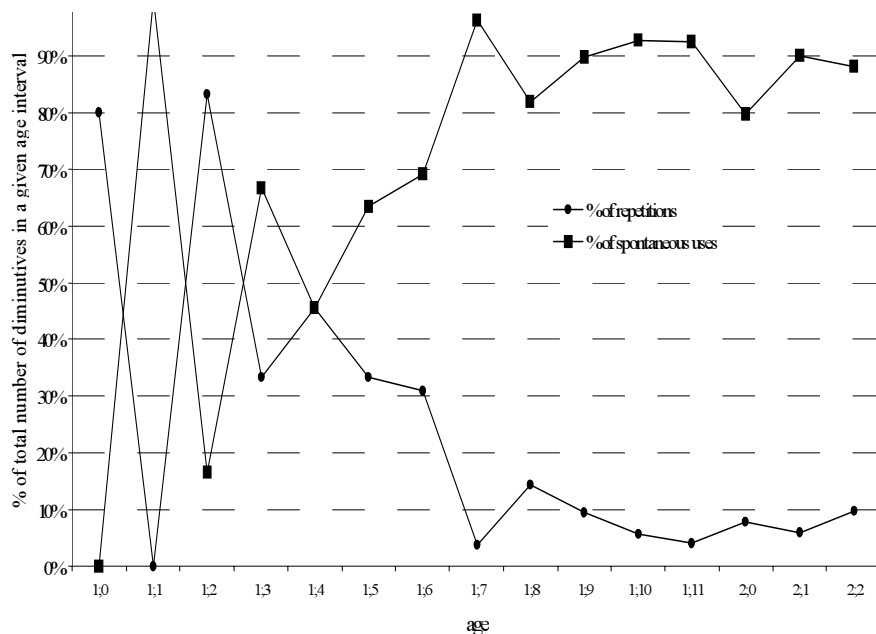
Thus all words which were not used in a spontaneous context at all were counted as repetitions. Figure 3. shows proportions of repetitions and spontaneous uses in total number of DIMs across age. It is evident that in the beginning (first five intervals – 1;0-1;4) the distribution of proportions is random. From the age 1;4 a systematic increase of spontaneous uses can be observed. The increase is statistically significant, measured by correlation with age for 11 age intervals (N=11, r=0.69, p<0.02).

Thus the systematic and significant increase of spontaneous uses of DIMs starting at age 1;5 particularly contrasts with the random distribution in the earlier period and seems to be a good indicator of productivity of the DIM category in Inka's speech.

Number of different formants used in DIM words

The number of different formants used at each age interval was counted. All phonetic and gender alternations were calculated separately, but the formant was counted once independently of the number of tokens and inflectional cases that occurred in the file. For example formants *-ka, -ko, -ek* (feminine, masculine and neuter gender of the formant *-k-*) were counted as 3 formants, but *–ki, -kę, -ką* (oblique cases of *-ka*) were counted as one formant. The reason for this was that at the beginning of the analyzed period Inka apparently hadn't established the category of grammatical gender. Thus one couldn't expect that she had established any connection among the different forms of the formant as one formant (*-k-*) in adult language. She practically never used any oblique cases (in the first several files) appropriately. (There were some systematic

Table 5. Summed frequencies of DIMs, number of formants and number of stems in Inka's speech at three age intervals

	1;0-1;4	1;5-1;9	1;10-2;2
DIM type frequency	30	334	641
Summed number of formants	17	64	96
Summed number of stems	28	236	350

"mistakes" where the oblique case form was used as nominative, but this was the only form of the word which Inka used in a given file). So initially there were no oblique case forms of formants to be calculated. Later, when Inka started to use oblique case forms appropriately it was evident that she distinguished inflection from derivation. Therefore oblique case forms of the formant were not counted separately.

The number of formants increases with age (exact data in Table 2.). The change is statistically significant (measured by correlation with age: N=15, r=0.86, p<0.00004). Hence one can conclude that Inka gradually incorporated into her lexicon DIM words with various formants. She not only increased the number of diminutives (as shown above) but also differentiated the set of formants used.

Another measure of formant productivity is shown on Fig. 4. Here, the number of formants is related to the number of diminutive types in each file. It can be seen that the proportion of formants to the total number of diminutive types in each file gradually decreases with age. Again a chi^2-test was conducted at three five-month age intervals. Table 5. shows the cumulated data of the number of formants and DIM types in Inka's speech.

The changes turned out to be statistically significant (chi^2=19.17; df=2; p<0.00001). Thus it may be stated that although both categories increase with age, the proportion of the two decreases significantly with age. That means that the more formants Inka used the more productive they were – she used not only more formants but used them for a greater number of DIM types. An additional chi^2 test for 2x2 tables comparing data in consecutive pairs of age intervals showed that the difference between the two first intervals is statistically significant (chi^2=11.4, df=1, p<.001), while between the second and third interval is nonsignificant (chi^2=1.99, df=1, p<.16). This again points to age 1;5 as an important developmental moment for the acquisition of DIMs.

Number of different stems used by Inka in DIMs

The next indicator of productivity of the diminutive category used in the analysis of Inka's diary was the number of different stems she used in DIM words. Stems were counted in one-month files. Each stem was counted as 1, independently of the number of different DIM word types in which it was used and independently of the number of formants attached to it. Data in Table 2 show the gradual increase of stem number with age. The increase measured by correlation with age is statistically significant (N=15; r=0.86; p<0.00004). Again (similarly to formants) the proportion of stems to the total number of DIM words at each age interval was computed. The results are shown on

Figure 4. Ratio of formants and stems to all diminutives in Inka's speech
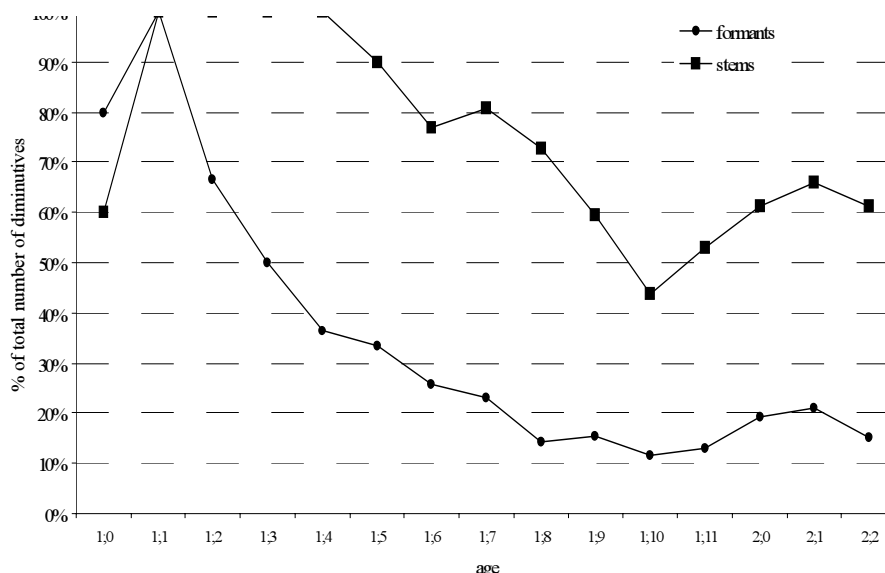

Fig. 4. In four of the five first files the number of roots equals the number of DIM types, which means that each root was used for one DIM type only. However, from age 1;5 systematic decrease of proportion can be observed (significant, tested by chi^2-test – chi^2=8.4; df=2; p<0.014, for three five-month age intervals – exact data in Table 5). This means that although both the number of stems used in DIMs and number of DIMs increase with age, the latter growth is larger – using more stems in DIMs results in proportionally a larger increase of DIM types. The effect is illustrated in Figure 5. Once more this indicates that from some point of development Inka used the DIM category productively.

The comparison of differences in the intensity of quantitative changes in stems, formants and DIM types used by Inka is illustrated on Figure 5. During the first five intervals the number of formants, stems and DIM types are almost identical and do not increase with age. When the number of formants rises to 10 (age 1;5) the small but systematic increase of stems and DIM types can be observed. When the number of formants exceeds 20 (age 1;9) the increase of stems and DIM types is already unproportionally larger. During the next two files (1;10 – 1;11) the number of stems and formants do not change considerably, but there is still an important growth of DIM types. This means that Inka massively combined stems and formants into new DIM types, which exactly is what one expects about a productive category.

Proportion of diminutives for which simple forms were found

The number of diminutives for which simple counterparts were found in Inka's speech was the last indicator of productivity. Only these simplicia, which were found in the same file where the DIM counterparts were located were taken into account (it

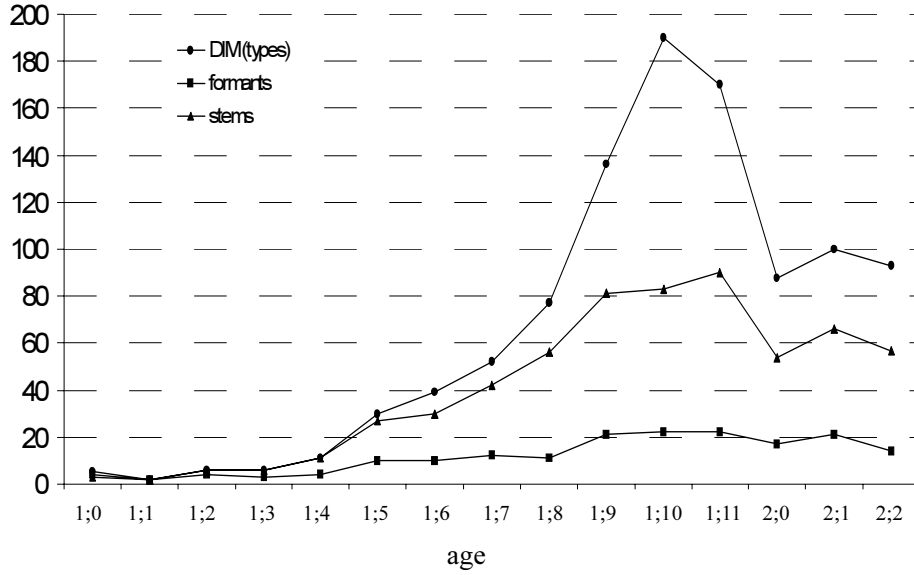Figure 5. Changes in DIMs frequency, formant and stem frequency in Inka's speech
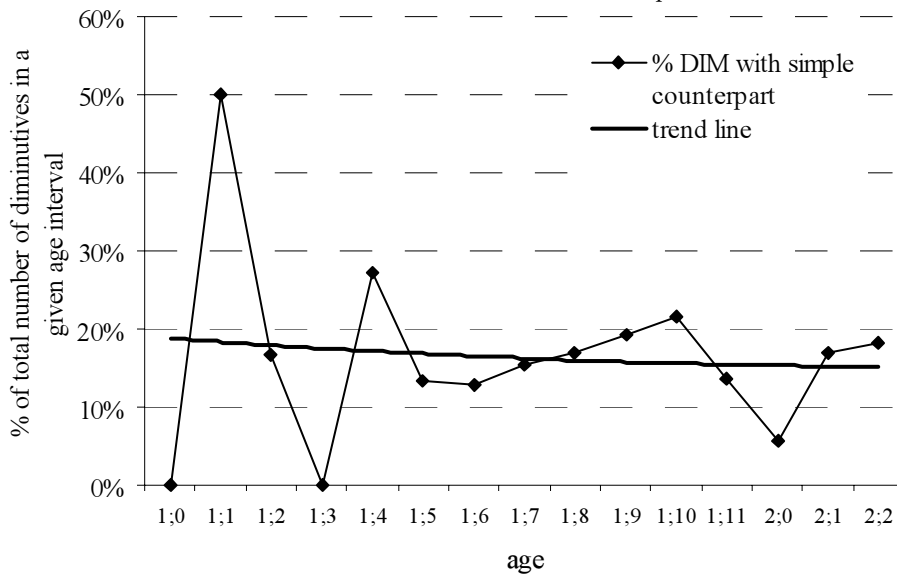

Figure 6. Proportion of DIMs for which a derivational base were found to total number of DIMs in Inka's speech

was not checked whether the simplex was present in earlier files). Although the number of diminutives for which simple forms are present in Inka's speech significantly increases with age (N=15, r= 0.66, p<0.008), there were no systematic changes in proportion of the number of "diminutives-with-simplex" to total number of DIMs across age (Pearson correlation for proportion of DIM-with-simplex in all DIM types is non-significant: N=15, r=-.28, p>.1; see: Figure 6).

The gradual increase of proportion with age could be taken as indirect evidence that DIMs are not used as simple – unanalyzed – words. It is probably much easier for a child to recognize the connection between a simple word and DIM derived from it if both are present in her/his speech. However this indicator unexpectedly gives no information about the productivity of the DIM category in Inka's speech. Several explanations are possible – for example: the lack of simplex can be due to the method of data collecting; paradoxically, a productive DIM form can replace its simple counterpart; finally, simple forms absent in production still can be comprehended.

When do diminutives start to be a productive category in Inka's speech?

Four of the five productivity indicators presented here clearly showed how the diminutive category becomes productive in Inka's speech during the analyzed period. Significant developmental changes not only in frequency, but particularly in proportion of various indicators to the total number of DIMs (or total number of words used in the case of the first indicator) showed that during the analyzed period the DIM category starts to be used productively. But is it possible to find a particular moment when Inka began to create DIMs as complex, derived words, and not as unanalyzed units? The comparison of changes in the four indicators suggests that indeed it is possible. In all four cases important changes start at the same age interval – at the age 1;5 (Figures 2, 3, 4, 5). This agreement in the starting point of changes in the frequency of four indicators is the additional argument for the productivity of the DIM category. To be certain that the observed differences are due to Inka's growing knowledge of derivational morphology and do not simply reflect characteristics of CDS, the last analysis of speech diary data was conducted. It was the analysis of proportions of DIMs to total number of words in CDS, which is described in the following section.

Proportion of DIM words in Child-directed Speech

No particular changes in DIM frequency in CDS should be expected in the analyzed period. In general Polish parents use a lot of DIMs when talking to small children, but there is no reason to expect a pattern of changes similar to that already observed in Inka's speech (Figure 2). Other potential pattern of changes in CDS could be a systematic decrease of DIMs. Adults usually adapt their CDS to the age of the child, and this rather means decrease of DIMs in CDS with the age of the child. However, in the case of Inka's early speech diary, it seems she was still too young at the end of the analyzed period to expect a lowering of DIMs in CDS.

The method of analysis of CDS in Inka's speech diary (1;0-2;02) was analogous to the analysis of Inka's speech (section 4). All DIMs were chosen from frequency lists made for all speakers present in Inka's diary except for herself. Table 6. presents

Table 6. Number of words used in CDS (types and tokens); type and token frequency of DIMs in CDS

Inka's age	1;0	1;1	1;2	1;3	1;4	1;5	1;6	1;7	1;8	1;9	1;10	1;11	2;0	2;1	2;2
CDS file size (types)	633	574	558	323	602	998	698	521	554	434	403	435	252	325	425
CDS file size (tokens)	2114	2118	2102	822	2276	4549	2755	1801	1714	1274	959	1179	538	708	1012
DIMs in CDS (types)	103	112	121	71	110	180	144	83	109	77	68	89	49	52	67
DIMs in CDS (tokens)	372	433	482	187	458	839	534	310	424	278	193	250	100	119	116

Figure 7. Proportion of diminutives to total number of words in CDS in Inka's diary (types and tokens)
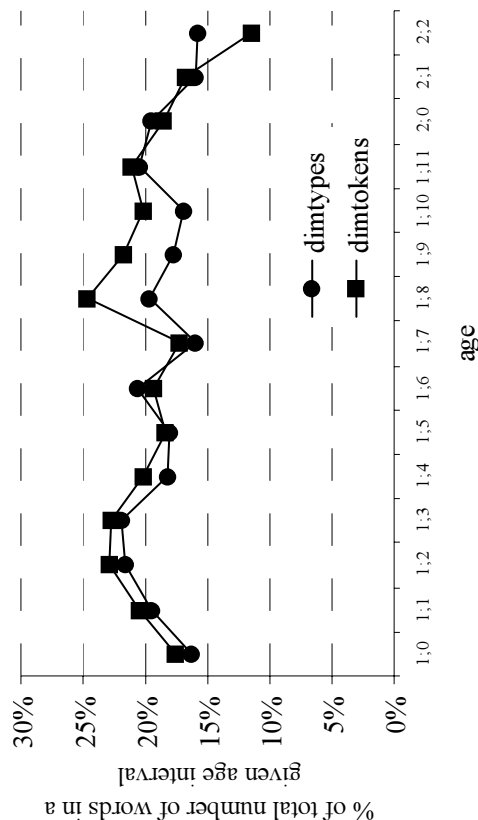
Table 7. Summed frequencies of DIMs and all other word types in CDS at three age intervals

	1;0-1;4	1;5-1;9	1;10-2;2
DIM type frequency	517	593	325
File size – DIM type frequency	2173	2612	1515

exact frequencies of DIMs in CDS both in types and tokens together with the file size (number of all words used in CDS in subsequent age intervals).

Figure 7 presents the proportion of DIMs to the total number of words in CDS at each monthly age interval. It is clearly visible that there are no substantial systematic changes in DIMs proportion in CDS. To test this assumption, a chi^2 test comparing DIM frequency with all other word frequencies (file size minus DIMs in types) was run on data cumulated into three 5-month age intervals (1;0-1;4, 1;5-1;9 and 1;10-2;2). The summed data used for a chi^2 test are presented in Table 7.

Indeed it occurred that there are no significant differences in the distribution of DIM type frequency in CDS (df=2, chi^2=1,76, p=.42).

For a statistical comparison of DIM frequency in Inka's speech (Table 4, Figure 2) and in CDS (Table 7, Figure 7), the data (type frequency of DIMs in Inka's speech and in CDS) were again cumulated into three age intervals. Chi^2 test (for 3x2 table) computed on these data turned out to be highly significant (chi2=550.6; p<.000001; df=2). Thus developmental changes ongoing in the distribution of DIM frequency across age in Inka's speech do not reflect the use of DIMs in CDS. This is an additional argument for the claim that the category of diminutives become productives in Inka's speech during the analyzed period. The proportion of DIMs (in types) in Inka's speech in lower than in CDS in the first five files (Figures 2 and 7), then rises and gains the level of DIMs in CDS in two consecutive files (age 1;5, 1;6) at about 18-20% and then in the next eight files (1;7-2;2) it is higher than proportion of DIMs in CDS on average 6.5%. Therefore the age 1;5 which was postulated as a starting point of productive use of DIMs by Inka is exactly the age at which she starts to use relatively the same amount of DIMs that adults use in CDS.

Discussion and conclusions

The quantitative analysis of changes in the early lexicon of a Polish child reveals that diminutives starts to constitute a productive derivational category from the age of 1;5. At this age the proportion of DIMs in the child's speech gains the level of DIMs in CDS and then exceed the proportion of DIMs in CDS on average 6.5%. At the same time the child starts to combine massively stems and formants in DIM words which results in a rapid enlargement of DIM types in her speech. All this suggest that the child has gained insight into the structure of the category of diminutives. Usage of various and numerous formants (the maximum of 22 formants at the ages 1;10 and 1;11) at the moment of most frequent DIM type production (190 and 170 DIM types respectively – Table 2) suggest that the productivity concerns the whole broad cat-

egory of diminutives, not particular formants. The present paper concerns the emer-
gence of word structure knowledge and does not raise the problem why diminutives
start to be a productive derivational category so early. The function(s) of DIM words
were not analyzed here, but it seems that the child has to acquire the structure first and
only then can use it in a certain function. This certainly does not mean that structures
are acquired independently of their function, but rather that the analysis of the acqui-
sitional process should include an analysis of structure productivity before asking
about its function (unproductive category cannot serve any functions at all). The func-
tion of diminutives in adult language is to depict the small size of an object or a posi-
tive emotional attitude (Grzegorczykowa, Puzynina, 1998). However, in languages
which make frequent use of the category (especially in CDS) diminutives are often
used much more frequently than the actual context of 'smallness' or 'niceness' would
account for. Thus it is not clear when and how children come to realize semantic
functions of diminutives existing in adult-adult speech. There are two main functions
of the category of diminutives proposed in the literature as reason for its early emer-
gence in child language. One is pragmatic – fictive function proposed by Dressler and
Merlini-Barbaresi (1999), who claim that diminutives in adult – child interactions are
used to emphasize non-seriousness of the situation. This idea was however criticized
by showing that in very first uses of diminutives (in various languages) it is very
difficult to identify any semantic or pragmatic function at all (Gillis, 1999). Gillis
claims that the very first uses of DIMs seem to be 'the play with pure form' rather than
to serve any pragmatic function. On the other hand, some researchers showed that in
languages which exploit diminutives greatly in CDS, the use of DIM forms (instead of
simple words) can simplify the inflection system or facilitate acquisition of other as-
pects of morphology, like grammatical gender (Kempe et al., 2001). Kempe et al.
demonstrated this on the example of Russian and Spanish, but certainly it can be the
Polish case as well. The problem is that efficient use of the DIM category for simpli-
fication of any aspects of inflectional morphology requires productivity of the cat-
egory. However, only 'structural' productivity is needed for this function. The data
used in Kempe et al. analysis came from speech diaries of children at the age 1;10-2;4.
Thus it is possible that they showed the consequences of DIM category acquisition
and not its emergence. The present analysis was aimed at showing the process of
emergence of this kind of knowledge. Neither semantic nor pragmatic functions of
diminutives were analyzed here, but a preliminary investigation of DIM contexts check-
ing (here described only in terms of spontaneous use of DIMs – section 4.2) suggests
that there is no evidence in Inka's diary (for the period 1;0-2;02) that the girl had
recognized either semantic or pragmatic functions of DIMs. For example, only a few
words were used consistently in a simple form when designating 'adult'/big object
(e.g. *auto* – a car) and in DIM form when designating 'child'/small object – like a
plaything (*autko* – a car DIM). In a great many contexts DIM words were freely used
for 'normal size' objects. This could favor the pragmatic – fictive function of DIMs,
but operationalization of this variable according to Dressler and Merlini-Barbaresi is
not clear. 'Non-serious' use of DIMs in adult-child interactions understood very broadly
as 'usage with regard to a child' reveals hesitation whether it is at all possible to

differentiate contexts where DIMs are used in this function from contexts where they are not, either by an adult or by a child.

In spite of the discussion on the potential functions of DIMs in early child language, it is important to consider the process of the development of the category productivity. The indicators of productivity adopted here demonstrate that even with lack of innovations it is possible to show how and when the derivational category of diminutives becomes productive in child speech.

References

Aitchison, J.(1994). *Words in the mind. An introduction to the mental lexicon*. Oxford: Blackwell.

Berko, J. (1958). The child's learning of English morphology, *Word*, 14, 150-77.

Bloom, P. (2000). *How children learn the meaning of words*. Cambridge, MA: MIT Press.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Chmura-Klekotowa, M. (1971). Neologizmy słowotwórcze w mowie dzieci. [Morphological neologisms in the speech of children.] *Prace Filologiczne*, 21, 99-235.

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge: Cambridge University Press.

Clark, E. V. (1995). Later lexical development and word formation. In: P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language*. Cambridge: Blackwell.

Czukowski, K. (1968). *Od dwóch do pięciu* [*From two to five*]. Warszawa:

Derwing, B. L., Baker, W. J.(1986). Assessing morphological development. In: P. Fletcher, M. Garman (eds) *Language acquisition*. 326–338.

Dressler, W. U., Merlini–Barbaresi, L. (1999). Morphopragmatics. In: J. Verschueren et al. (Eds.) *Handbook of Pragmatics*. Amsterdam: Benjamins. 1-14.

Gillis, S. (1999). Paper presented at VIIIth International Congress for the Study of Child Language, 12-16 July, San Sebastian.

Grzegorczykowa, R., Puzynina, J. (1998a). Problemy ogólne słowotwórstwa [General problems of word formation]. In: R. Grzegorczykowa, R. Laskowski, H. Wróbel (Eds.), *Gramatyka współczesnego języka polskiego. Morfologia*. [*Grammar of the modern Polish language. Morphology*.], 361-387.Warsaw: PWN.

Grzegorczykowa, R., Puzynina, J. (1998b). Rzeczownik. [Noun.] In: R. Grzegorczykowa, R. Laskowski, H. Wróbel (Eds.), *Gramatyka współczesnego języka polskiego. Morfologia* [*Grammar of the modern Polish language. Morphology*.], 389-464. Warszawa: PWN.

Haman, E. (2000). Status wyrazów pochodnych w języku dziecka. [The status of complex words in a child language]. Unpublished doctoral dissertation. University of Warsaw.

Haman, E. (2002). Derived words in the lexicon of Polish children: The role of word structure in derivational rule acquisition. *Annual Review of Language Acquisition*, 2, 59-84.

Kempe, V.,Brooks, P.J., Pirott, L. (2001). How Can Child-Directed Speech Facilitate the Acquisition of Morphology? In: M. Almgren, A. Barreńa, M.J. Ezeizabarrena, I. Idiazabal, B. MacWhinney (Eds.), *Research on Child Language Acquisition. Proceedings of the 8th Conference of the International Association for the Study of Child Language.* Somerville: Cascadilla Press.

MacWhinney, B. (1995). *The CHILDES project: tools for analyzing talk. Hillsdale*, NJ: Lawrence Erlbaum Associates.

Nagórko, A. (1998). *Zarys gramatyki polskiej (ze słowotwórstwem).* [*An outline of Polish grammar and word formation*]. Warszawa: PWN.

Savickiene, I. (1998). The acquisition of diminutives in Lithuanian. In: S. Gillis (Ed.), *Studies in the acquisition of number and diminutive markin*g, Antwerp Papers in Linguistics 95, Antwerp: Universiteit Anwerpen, 115-135.

Smoczyńska, M. (1998). Krakowskie dane językowe dzieci w systemie CHILDES. [Cracow child language data in CHILDES system]. In: Smoczyńska, M. (Ed.), *Studies in Developmental Psychology and Psycholinguistics*. Cracow: Universitas.

Tomasello, M. (1992). *First verbs: A case study in early grammatical development*. Cambridge, MA: Cambridge University Press.