White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/1587/

Article:
Dabrowska, E. (2001) Learning a morphological system without a default: the Polish genitive. Journal of Child Language, 28 (3). pp. 545-574. ISSN: 0305-0009

https://doi.org/10.1017/S0305000901004767

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

J. Child Lang. **28** (2001), 545–574. © 2001 Cambridge University Press
DOI: 10.1017/S0305000901004767 Printed in the United Kingdom

Learning a morphological system without a default: the Polish genitive*

EWA DĄBROWSKA
University of Sheffield

(*Received* 4 *November* 1999. *Revised* 28 *September* 2000)

ABSTRACT

The acquisition of the English past tense inflection is the paradigm example of rule learning in the child language literature and has become something of a test case for theories of language development. This is unfortunate, as the idiosyncratic properties of the English system of marking tense make it a rather unrepresentative example of morphological development. In this paper, I contrast this familiar inflection with a much more complex morphological subsystem, the Polish genitive. The genitive case has three different markers, each restricted to a different subset of nouns, in both the singular and the plural. Analysis of the spontanous speech of three children between the ages of 1;4 and 4;11 showed that they generalized, and overgeneralized, all three singular endings. However, error rates were extremely low and there is no evidence that they treated any one ending as the 'default'. The genitive plural, on the other hand, showed a strikingly different pattern of acquisition, similar to that seen in English-speaking children learning the past tense. It is argued that in the latter two cases, the default-like character of one of the affixes is attributable to the properties of the relevant inflectional subsystems, not to the predispositions that children bring to the language-learning task.

INTRODUCTION: THE RULE DEBATE

For most linguists, it is axiomatic that lexical and grammatical knowledge are distinct aspects of a speaker's competence. The lexicon is finite, whereas the number of complex units (inflected forms, phrases, and sentences) that a speaker can produce and understand is in principle infinite. Furthermore, the relationship between a word and its referent is arbitrary and hence un-

[*] I would like to thank Magdalena Smoczyńska for undertaking the Herculean task of tagging the Polish corpus and for making the data available to me, and Ann Peters and two anonymous referees from JCL for their comments on earlier versions of this paper. Address for correspondence: Dr. Ewa Dąbrowska, Department of English Language and Linguistics, University of Sheffield, Sheffield S10 2TN, UK.

predictable, while complex units can be derived by combining simple units according to a finite set of rules. Clearly, words must be stored in memory, while complex units are assembled by some mental analogue of a formal grammar.

It follows that language acquisition involves learning words and the rules for inflecting and combining them. No one doubts, of course, that children learn words; and the child language literature offers ample evidence that they are able to construct forms which they could not have heard from their parents. Jean Berko's famous 'wug' experiment (Berko, 1958), replicated by a number of other researchers, clearly showed that children are able to inflect novel words. In naturalistic settings, virtually all children occasionally produce regularizations such as **breaked* and **foots*, often after a period during which they had supplied the correct irregular. Until the mid-eighties, nearly everyone agreed that these well-known facts provided direct evidence that children learn symbolic rules of the kind postulated by linguists.

This seemingly unshakeable view has been challenged by the connectionists. In 1986, Rumelhart & McClelland reported that they had built an artificial neural network which had successfully learned to produce past tense forms of English verbs. They had trained the network by presenting it with pairs of verb stems and past tense forms, and tested it on verbs in the original training set as well as on novel verbs. The network was able to provide the correct past tense forms of many (though not all) of the novel verbs. Moreover, it sometimes regularized irregular verbs it had learned earlier, thus exhibiting the U-shaped development observed in children learning the past tense. Thus, in at least two important respects, the model's performance resembled that of real children. What was particularly interesting from a psychological perspective was that the network represented and processed both regular and irregular verb forms using the same mechanism: connection weights. This suggested that humans, too, could use the same mechanism – neural connections of varying strengths – when dealing with both kinds of verbs. Furthermore, the network exhibited rule-like behaviour without an explicit representation of any 'rules', which seemed to accord well with the tacit nature of linguistic knowledge.

As critics were quick to point out, there were many problems with this early model (Pinker & Prince, 1988). Later models addressed some of these problems with considerable success (see, for example, Plunkett & Marchman, 1993; Hare, Elman & Daugherty, 1995), forcing proponents of what came to be known as the dual mechanism theory to acknowledge that connectionist nets can model some aspects of human linguistic knowledge and that they are capable of behaving productively, i.e. generalizing a previously learned pattern to novel input (cf. Pinker & Prince, 1992; Marcus, Pinker, Ullman, Hollander, Rosen & Xu, 1992; Marcus, Brinkman, Clahsen, Wiese & Pinker, 1995; Pinker, 1998). However, they argue that the productive behaviour

exhibited by these models is fairly limited and corresponds to the kind of productivity that is characteristic of IRREGULAR inflections in human language. They are adamant that this process is different in kind from that which underlies regular inflections.

Proponents of the dual mechanism theory point out that irregular inflections are typically restricted either to single words (*go-went*) or to small clusters of phonologically similar words (*sink-sank*, *drink-drank*, *stink-stank*). Regular inflections, on the other hand, apply to all kinds of stems, regardless of their phonological properties. Since connectionist models operate by comparing a stimulus to stored representations, they perform very well on irregular words. However, for exactly the same reason, they have great difficulty in learning to apply the regular inflection to novel stimuli which do not resemble previously learned exemplars. Rumelhart & McClelland and their followers have been able to overcome this problem by ensuring that a substantial majority of the verbs on which their models were trained were regular. In such circumstances, with some tinkering with the parameters, nets can be coaxed into supplying the regular inflection most of the time, although they still tend to perform better on novel forms which resemble previously learned verbs.

However, according to Clahsen, Rothweiler, Woest & Marcus (1992), Marcus *et al*. (1995), and Pinker (1998, 1999) regularity in the psychological sense does not depend on frequency: humans learn the regular inflection, and extend it to nearly all novel words, even if it is rare in the input. To explain this ability, we must appeal to a different mechanism: symbolic rules. Symbolic rules operate by adding an affix (e.g. *-ed*) to an abstract mental symbol representing a class of words (e.g. verbs), and therefore are not constrained by similarity to stored exemplars. The human language processing system, then, uses two mechanisms: symbolic rules for regular processes and associative memory for irregular processes.

Proponents of the dual mechanism theory support this position with an impressive array of arguments. They point out that irregular inflections rely on memory, and hence can be applied only when memory can be accessed. Thus, the theory predicts that when the inflected form cannot be retrieved from memory, either because it is not there or because it is inaccessible, speakers would have to use the regular or 'default' mechanism. And indeed, in a variety of circumstances when access to memory is ruled out, speakers do resort to the regular inflection – in English, the *-ed* ending for the past tense and *-s* for plural. Marcus *et al*. (1995) enumerate 21 such circumstances. The first sixteen of these, listed in Table 1, are various categories of words which require the default ending: that is to say, if another inflection were used, the resulting structure would be ill-formed. The remaining five are rather different in that they involve errors – i.e. ungrammatical forms – produced by various populations of speakers (normal children and

TABLE 1. *Circumstances in which memory is not accessed and the regular inflection is applied* (*based on Marcus* et al. *1995*)

Circumstance	Examples
1. Novel words	*snarfed*, *wugs*
2. Low-frequency words	*stinted*, *eked*
3. Unusual-sounding words	*ploamphed*, *krilged*
4. Words with irregular homophones	*lied/lay*, *hanged/hung*
5. Words which rhyme with irregular words	*blinked*, *glowed*
6. Onomatopoeic words	*dinged*, *peeped*
7. The word is mentioned rather than used	*I checked the article for sexist writing and found three* mans *on page 1.*
8. Surnames	*the Childs*, *the Manns*
9. Unassimilated borrowings	*latkes*, *cappucinos*
10. Truncations	*synched*, *mans* 'manuals'
11. Acronyms	*PACs*, *OXes*
12. Derivation from different category:	
(a) denominal verbs	*spitted* 'put on a spit'
(b) deadjectival verbs	*righted* 'returned to upright position'
(c) nominalizations	*ifs*, *ands*, *buts*
13. Derivation via different category:	
(a) via noun (V → N → V)	*costed* 'calculated the costs'
(b) via verb (N → V → N)	*wolfs* 'instances of wolfing'
14. Derivation via name	*Mickey Mouses*, *Renault Elfs*
15. Bahuvrihi compounds	*sabre-tooths*, *low-lifes*
16. Nominalized phrases	*bag-a-leafs*, *shear-a-sheeps*

adults, Alzheimer's patients, anomic aphasics, and Williams syndrome sufferers). These are discussed separately here.

It is not difficult to see that the list in Table 1 is an odd assortment of seemingly unconnected and sometimes rather exotic circumstances. The very arbitrariness of this collection is a strong argument in favour of the dual mechanism theory: without the concept of a default inflection, it would be difficult to explain why the same ending should be used in all these contexts, not just in English, but in other languages as well (see below).

There is also a fair amount of psycholinguistic and neurolinguistic evidence supporting the psychological reality of the distinction between regular and irregular verbs. First, although both children and adults occasionally overgeneralize the *-ed* ending (producing regularizations such as *comed* or *bringed*) and the various irregular patterns (producing irregularization errors such as *bat* for *bit* and *truck* for *tricked*), the latter kind of errors is much less frequent (Marcus *et al.*, 1992; Xu & Pinker, 1995). The largest study of the overgeneralization of the regular past tense ending by English-speaking children, Marcus *et al.*, 1992, reports a mean regularization rate of 4·2% (median 2·5%, range 0–24%). An earlier study by Kuczaj (1977) reports a much higher incidence of regularization errors (mean

20·9%, range 1·1%–40·2%), but Marcus and his co-authors argue that these figures are inflated by Kuczaj's elicitation procedure. Marcus *et al*. also calculated regularization rates for children in Miller & Ervin's and Valian's samples; the relevant figures are 10·2% and 5·1%, respectively. In contrast, according to Xu & Pinker (1995), English-speaking children irregularize past tense forms only 0·19% of the time. Thus, regularization rates for the English past tense are at least 22 times, and possibly as much as 110 times higher than irregularization rates.

Secondly, several experimental studies show that performance on irregular verbs is highly sensitive to frequency and phonological similarity to other irregulars: subjects take longer to react to low frequency irregulars than to irregular verbs of higher frequency, and judge irregular past tense forms of novel verbs as more acceptable if they resemble existing irregulars. Regular verbs, in contrast, are claimed not to show such effects (Pinker & Prince, 1992; but see Marchman, 1997).

Finally, there is evidence suggesting that the ability to produce regular and irregular inflections may be differentially impaired in various neurological disorders. In some clinical populations (e.g. anomic aphasics, Alzheimer's patients), the ability to produce irregular inflections is impaired while regular processes appear to be unaffected. Conversely, in agrammatic aphasics, Parkinson's patients, and children with Specific Language Impairment, the grammatical system may be severely disrupted while word retrieval, including the retrieval of irregular forms, remains comparatively good (Pinker, 1998, 1999; Pinker & Prince, 1992; Marcus *et al*., 1995). This suggests that different neural systems may be involved in processing regular and irregular inflections; and indeed some neural imaging studies on normal subjects (e.g. Jaeger, Lockwood, Kemmerer, Van Valin & Murphy, 1996) appear to support this claim, although their conclusions have not gone unchallenged (see Seidenberg & Hoeffner, 1998).

THE PSYCHOLINGUIST'S FRUITFLY?

The English past tense has been a favourite object of study for both the connectionists, who have repeatedly tried to model its acquisition, and for proponents of the dual mechanism theory, who found in it a rich source of evidence for dissociations between the regular and irregular inflections. In this way, this rather unassuming inflection has become something of a test case for theories of language acquisition and processing. As Pinker observes,

> 'The recent flurry of studies on the neurology of the past tense – aside from its contribution to the connectionism debate – may offer hope for a better understanding of language and the brain in general. Irregular and regular verbs are nicely matched in complexity and meaning; and regular inflection, which people compute so freely when faced with new verbs, is

> perhaps the simplest example of the great human capacity for generating an unlimited number of new linguistic forms. Perhaps regular verbs can become the fruitflies of the science of language – their recombining units are easy to extract and visualize, and they are well studied, small, and easy to breed.' (Pinker, 1997: 548)

This 'flurry of studies' has without a doubt added to our knowledge of the psychology and neurology of the English past tense, and has given us a much better understanding of the potential and the shortcomings of artificial neural networks. Whether it can fulfil Pinker's hopes is less certain, and depends to a large extent on whether conclusions about English past-tense morphology can be generalized to other morphological systems, and linguistic rules in general. Unfortunately, there are reasons to suspect that they cannot.

The problem is that the acquisition of the English past tense is not a very representative example of morphological learning. To begin with, the past tense rule is extremely simple. This, of course, is one of the reasons why it became a favourite object of study in the first place, but its very simplicity makes it unrepresentative. More seriously, in the English tense-marking system, regularity is inextricably bound up with several other properties, which makes it difficult to determine whether any observed differences between regular and irregular verbs are due to regularity *per se* or to one or more of these contingent factors.

First, the regular and irregular inflections invoke different morphological mechanisms to mark tense. With regular verbs, past tense is marked by suffixation, resulting in a form which is easily analysable into a component which specifies the type of activity (e.g. *play*) and a component which indicates its time relative to the speech event (*-ed*). Most irregular verbs require vowel changes – a much less transparent, and hence less readily generalizable, method of marking the same distinction. This asymmetry may be partly responsible for the fact that regularization errors (e.g. *choosed* for *chose*) are vastly more frequent than irregularization errors (e.g. *snoze* for *snoozed*). A regularization error may, as suggested by proponents of the dual mechanism theory, occur when the child is unable to retrieve the correct irregular from memory; but it could also arise from an attempt to signal the past tense more clearly. In most irregular forms the tense marker is, in effect, buried inside the stem. Thus, a child who wants to emphasize the fact that an action occurred in the past might choose to use the regular form even if he/she knows the correct irregular, rather like an adult who says *pig meat* instead of *pork* in order to draw attention to the connection between a sizzling piece of bacon and the mud-wallowing animal. This explanation is particularly plausible in the case of double-marked forms such as *camed* or *sanged*, which are unlikely to result from retrieval failure.

Secondly, many English irregular verbs form phonologically similar

clusters (e.g. *ring-rang*, *sing-sang*, *spring-sprang*), while the regular inflection applies to verbs irrespective of their phonological properties. It is often assumed that this is true more or less by definition, as in the following passage:

> 'Irregular forms belonging to the same class always share many of their phonological (and/or other) properties and one can therefore propose that they can always be plotted in a coherent area of the input space (given appropriate coding). In contrast, forms falling into the regular class are usually very heterogeneous and they would, therefore, occupy different and even discontinuous parts of the input space.' (Pulvermüller, 1998: R31)

However, even the most cursory survey of the inflection systems of the world's languages reveals that this is simply not the case. In languages with morphological classes, 'regular' inflections are restricted to forms sharing certain phonological or grammatical features. To take a well-known example, in Spanish (and other Romance languages), verbs are assigned to 'conjugations' on the basis of which vowel occurs after the verb root in certain forms. Verbs belonging to different conjugations require different endings to signal the same grammatical function: for instance, the first person singular preterite ending is *-é* for verbs belonging to the first conjugation (which end in *-ar* in the infinitive) and *-í* for verbs belonging to the second and third conjugations (which end in *-er* and *-ir* in the infinitive). Conversely, some languages have irregular classes which are phonologically (and semantically) arbitrary: what defines these classes is that they share a particular inflection or set of inflections. The class of masculine nouns that take the *-u* ending in the genitive singular in Polish is one example (see below).

Another complicating factor is frequency. Although both classes contain rare as well as highly frequent verbs, as a group, irregular verbs have very high token frequency: according to Marcus *et al.* (1992: 75), about 85% of past tense forms in speech addressed to children are irregular. Regular verbs, on the other hand, have very high type frequency: about 86% of the 1000 most frequent verbs in English (and 95% of the entire verb lexicon) take the *-ed* ending in the past tense. High token frequency clearly favours rote learning, since it is easy to memorize a frequently occurring form – in fact, as Bybee (1995) points out, very high token frequency may actually 'protect' a verb from being subsumed under a schema, thus preventing the extraction of shared patterns. High type frequency, on the other hand, facilitates schema extraction: it is easier to notice a pattern shared by three hundred verbs than one which is shared by only three. Consequently, the differences in frequency exaggerate the differences between regular and irregular verbs.

The regular past tense inflection in English is thus highly productive, both in the quantitative sense (it applies to the vast majority of verb types) and in

terms of applicability (it is phonologically unrestricted). What is more, it is also the ONLY truly productive past tense inflection in English: the dozen or so irregular patterns are either non-productive or only marginally productive, and apply to narrowly circumscribed groups of stems. In view of this fact, it is hardly surprising that the same inflection is used in each of the sixteen circumstances enumerated in Table 1: the 'default' ending is, in practice, the only inflection available for use with new words.

REFINING THE QUESTION: THE GERMAN PLURAL

It is clear, then, that the morphology of the English past tense is not a very good testing ground for the psychological reality or otherwise of the dual mechanism theory. Two other inflections which have been considered in this connection – the English plural (Marcus, 1995) and the German past participle (Marcus *et al.*, 1995) – suffer from similar drawbacks. However, there is a much more promising source of evidence which has recently attracted a great deal of attention: the German plural.

The plural in German can be marked by one of five affixes (*-n*, *-e*, *-er*, *-Ø*, *-s*), three of which (*-Ø*, *-e* and *-er*) are sometimes accompanied by vowel changes in the stem. Which affix is used with which noun is determined partly by gender and partly by morphophonological properties of the stem. However, one of the endings, *-s*, has a special status. It is less restricted phonologically than the other affixes, and it is used in various 'emergency' situations: with novel or unusual sounding words and with names and other noncanonical roots such as clippings, acronyms, onomatopoeic words, unassimilated borrowings, and conversions. Significantly, these are the very circumstances in which the default inflection applies in English (cf. Table 1). It would appear, then, that German does have a regular or 'default' plural ending – namely, *-s* (Clahsen *et al.*, 1992; Marcus *et al.*, 1995; Clahsen, 1999).

It is not difficult to see that the German plural makes a much better fruitfly than the English past tense. It uses the same morphological mechanism in both regular and irregular inflections, and the irregular endings are clearly productive, which makes possible a direct comparison of the kind of productivity exhibited by both patterns. Furthermore, the fact that the *-s* ending is fairly rare (it is used with only about 4% of German nouns) makes the German plural system the perfect test case for a central tenet of the dual mechanism theory: the claim that default status does not depend on frequency.

Before we proceed, it is worth noting that, although *-s* does tend to be preferred in 'default' circumstances, it is not the only affix that can occur in such contexts. Marcus *et al.* themselves note that bahuvirhi compounds in German behave just like ordinary endocentric compounds, that is to say, they take whatever ending is required by the right-most element. Another fairly

systematic exception is low-frequency words: if *-s* only applies to 4% of German nouns, it follows that the vast majority of low-frequency words must take other endings. The 'irregular' inflections are also sometimes used with nominalized verbs, nominalized VPs, names (especially place names and product names), acronyms, and many borrowings, including very recent borrowings (see Köpcke, 1988; Wegener 1994, 1999). Thus, although the German *-s* plural is indeed special, it does not have quite the same status as its English counterpart, or the English regular past tense. However, from our perspective, a more interesting question is whether GERMAN SPEAKERS treat *-s* as the default inflection and whether the psychological mechanisms underlying its use are indeed qualitatively different from those invoked by the irregular affixes. It is to these questions that we now turn.

One promising source of evidence bearing on the issue is experimental studies using novel words as stimuli. Because speakers do not have lexical entries for words which do not exist in the language, experimental tasks involving novel words should tap their knowledge of the default process. In one such study (Köpcke, 1988), adult subjects were asked to pluralize novel words. The results, however, were not what the dual mechanism theory would lead us to expect. Speakers chose different plural affixes, depending on properties of the stem: *-e* with masculine monosyllabic nouns, *-n* with monosyllabic feminines and feminines ending in schwa, and so on.

Marcus *et al.* (1995) argue that Köpcke's experiment is irrelevant to the rule debate because he presented nouns in isolation, and so subjects treated them as roots – and presumably analogized from known roots. In their opinion, to determine whether German speakers treat *-s* as the default plural affix, one needs to rule out access to memory for roots. (This is strange, since use with novel words tops their list of circumstances in which the default inflection is used – cf. Table 1.) In their own experiment, Marcus *et al.* presented novel words in one of three conditions: as roots (e.g. 'I have taken a green kach for my cold'), as names ('My friend Hans Kach and his wife Helga Kach are a bit strange'), or as borrowings ('The French ''kach'' looks best in black'). The subjects' task was to rate the naturalness of sentences containing the plural forms of these novel nouns on a scale from 1 to 5. The authors then compared the ratings for the *-s* plural and the highest-rated irregular.

Marcus and his colleagues found that when the novel nouns were presented as ordinary words (i.e. roots), subjects judged the irregular forms as more acceptable – a result that echoes Köpcke's data. However, when they were presented as names, the regular plurals received higher ratings; and when they were presented as borrowings, the regular form and the best irregular were judged to be equally good. Furthermore, irregular plurals tended to get higher ratings when the stimuli rhymed with ordinary German words; the regulars did not show this effect. (In fact, *-s* plurals tended to get

slightly higher ratings if they DID NOT rhyme with real words.) In sum, their results show that all the German plural affixes can be applied to novel nouns, though *-s* typically applies in somewhat different circumstances than the rest.

Another source of psycholinguistic evidence for the dual-mechanism theory comes from studies of overgeneralization errors in child language. According to the theory, the regular or 'default' inflection applies whenever the inflected form cannot be retrieved from the lexicon, either because it is not there or because it is inaccessible. Since children's lexical entries for irregular words are not yet well established, they are sometimes unable to retrieve the correct form from memory and apply the regular inflection instead. Thus, the dual mechanism theory predicts that German-speaking children will sometimes overgeneralize the *-s* ending – and indeed they do. The problem is that they overgeneralize ALL the plural endings, and the most commonly overgeneralized one is the highly frequent *-n*, not *-s* (Clahsen *et al.*, 1992; Köpcke, 1998).

Nevertheless, Clahsen *et al.* (1992) maintain that these findings provide support for the dual mechanism theory, arguing that most children overgeneralize *-n* because they THINK that it is the default ending. They feel that such an interpretation is justified by two striking differences between *-n* and *-s* on the one hand and the irregular *-er* and *-e* on the other.

First, they observe that 'the rates of [overgeneralization] for *-n* (22·5%) and *-s* (25%) are an order of magnitude higher than the rates for *-e* (2·5%) and *-er* (1·0%), suggesting that qualitatively different mechanisms are responsible for these [overgeneralizations]' (Clahsen *et al.*, 1992: 247–8; the authors actually use the term 'overregularization' not 'overgeneralization', but it is clear that the term is intended to cover both regularizations and irregularizations). These figures are derived from an analysis of the speech of their main pool of subjects – nineteen dysphasic children between the ages of 3;1 and 6;11. Clahsen *et al.* assure us that, as far as pluralization is concerned, 'there are no differences … between normal and dysphasic children' (241). This, however, does not seem to be borne out by their data, as their one normal control, Simone, does not show such dramatic differences in overgeneralization rates as the dysphasic children (see Table 2; Clahsen *et al.* have forgotten to include the overgeneralization rates for Simone in their

TABLE 2. *Relative overgeneralization rates for German plural affixes* (*Data from Clahsen* et al. *1992*)

	Tokens correct					Overgeneralizations				ROG rate (%)			
	-n	*-s*	*-e*	*-er*	*-Ø*	*-n*	*-s*	*-e*	*-er*	*-n*	*-s*	*-e*	*-er*
Simone	83	28	130	47	57	4	4	1	3	4·60	12·50	0·76	6·00
Dysphasic	224	39	214	104	132	65	13	6	1	22·49	25·00	2·73	0·95

paper, so the figures given here have been calculated on the basis of the data they provide in their Table 3).

Other studies of normal German-speaking children provide no evidence that overgeneralization errors involving *-n* and *-s* are 'an order of magnitude' more frequent that those involving the other two affixes. Köpcke (1998) summarizes unpublished data on seven normal children, collected by Clahsen and his colleagues (presumably after the publication of the 1992 paper). Between them, the children overgeneralized *-n* fifteen times, *-e* seven times, and *-s* only two times. Park (1978), Mills (1985), and Bittner & Köpcke (2000) report similar findings; and in an experimental study conducted by Mugdan, the most frequently overgeneralized affix was *-e* (1977, cited in Köpcke, 1998). These results suggest that we must be very careful in interpreting the data from the dysphasic children.

There is also another reason to be cautious about the conclusions put forward by Clahsen and his colleagues. Overgeneralization rates are usually calculated as a ratio of the number of times a particular affix has been overgeneralized to the number of opportunities for overgeneralizing it (i.e., the number of stems which require some other inflection):

(1) OG rate for affix X

$$= \frac{\text{tokens of OG of affix X}}{\text{tokens of OG of affix X} + \text{tokens correct with other affixes}} \times 100$$

Clahsen *et al*. use a different measure, which I will call the 'relative overgeneralization rate' or ROG. The ROG for a particular affix is the ratio of overgeneralization tokens to all tokens *with that affix*:

(2) ROG rate for affix X

$$= \frac{\text{tokens of OG of affix X}}{\text{tokens of OG of affix X} + \text{tokens correct with X}} \times 100$$

Clahsen *et al*. do not explain why they have decided to use this particular method of calculating overgeneralization rates.[1] This is unfortunate, as employing the usual method would give very different results: it would obliterate the differences between *-s* and the irregulars in the dysphasic group, and give very similar overgeneralization rates for all four affixes in the normal subject (see Table 3). Thus, it is simply not the case that *-s* is overgeneralized more frequently than the other affixes in any absolute sense, though the data in Table 2 show that it is overgeneralized more frequently than we would expect on the basis of frequency alone. But the only legitimate

[1] This is puzzling in view of the fact that Marcus is a co-author. In an earlier study (Marcus *et al*., 1992), the OG formula was used.

TABLE 3. *Overgeneralization rates for German plural affixes calculated as a proportion of opportunities for overgeneralization errors*

Subjects	Overgeneralization rate (%)			
	-n	*-s*	*-e*	*-er*
Simone	1·50	1·25	0·46	1·00
Dysphasic	11·73	1·89	1·19	0·16

conclusion that we can draw from this is that frequency is not the ONLY determinant of overgeneralization.

The second argument that Clahsen *et al*. use to support their claim that most of the children in their study misclassified *-n* as the default ending hinges on a curious fact about the use of plurals in compounds. In German, as in English, the regular plural does not occur inside compounds: thus, compounds such as **Auto-s-berg* 'car-heap' and **Sozi-s-treffen* 'socialists' meeting' are impossible, while *Büch-er-regal* 'book shelf' (with the plural *-er*) and *Frau-en-laden* 'women's centre' are fine (cf. English **rats-infested* v. *mice-infested*). This seemingly bizarre restriction can be given a simple explanation in the framework of Kiparsky's level-ordering morphology (Kiparsky, 1985). In Kiparsky's model, irregular inflections are formed at level 1, compounding occurs at level 2, and regular inflections are attached at level 3. In the course of a derivation, all level 1 processes apply before those at level 2, which in turn precede those at level 3. Because irregular morphological processes apply before compounding, irregular plurals can occur inside compounds; but regular inflections, which are added at level 3, cannot be attached to a constituent of a compound assembled at level 2, only to the whole compound. Thus, if a child misclassifies *-n* as the regular ending, it should apply at level 3, after compounding, which would preclude it from occurring inside compounds. Intriguingly, Clahsen *et al*. found that children who overgeneralized the *-n* plural often left it out in compounds, producing forms such as *Dose-öffner* 'can opener' instead of *Dose-n-öffner*, *Bauer-hof* 'farm' (lit. 'farmer-yard') instead of *Bauer-n-hof*, and *Küche-fenster* 'kitchen window' instead of *Küche-n-fenster*; but they never left out *-er* or *-e*. This, Clahsen and his co-authors argue, shows that children are sensitive to the distinction between regular and irregular inflections.

This explanation is interesting, but it is not without problems. The seven children who overgeneralized *-n* did not invariably leave it out in compounds: in fact, they actually supplied it, on average, 63% of the time. In contrast, English-speaking children almost never use the regular plural inside compounds, even novel compounds formed in an experimental setting (Gordon, 1985). Since German children hear the plural *-n* inside compounds in the

input, it is not particularly surprising that it appears in their own speech as well; but the fact that they use it most of the time is difficult to reconcile with Clahsen's argument. It is also worth noting that all four of the plural affixes are sometimes used as linking morphs inside compounds, where their function is to join together the elements of the compound rather than to signal plurality. For example, *Bauernhof* refers to something belonging to a single farmer, not farmers; and *Küchenfenster* means 'window in the kitchen' not 'window in the kitchens'. Although the children in the study also left out *-n* in genuinely plural compounds, about half of the compound-internal *-n*'s which Clahsen and his collaborators counted as plurals were in fact semantically singular. Finally, it is probably no coincidence that the two plural affixes that were never deleted inside compounds, namely *-er* and *-e*, are syllabic, while *-s* and *-n* are not.[2] Hence, the former are unlikely to be omitted for purely phonological reasons: leaving out a whole syllable would dramatically affect the overall rhythmic structure of the word.

To conclude: the German research does show that high frequency is not a necessary precondition for generalization. Moreover, the fact that the *-s* ending is preferred in most of the circumstances enumerated in Table 1 suggests that it does have a special status, and this certainly requires an explanation.[3] Nevertheless, there is little evidence that German speakers treat it as the default plural marker, since the other affixes are also used in circumstances which are said to call for the regular inflection, i.e. with novel words and when memory access fails, by both children and adults. To properly evaluate the dual mechanism theory, it is necessary to consider data from another language, or better still, from a number of languages. In the remainder of this paper, I will examine one specific example, the Polish genitive, which, as we shall see, has certain properties which make it a particularly interesting test case.

THE POLISH GENITIVE: A DESCRIPTION OF THE SYSTEM

Polish nominal inflections conflate two grammatical categories: case and number. There are seven cases (nominative, genitive, dative, accusative, instrumental, locative, and vocative) and two numbers (singular and plural). The inflectional endings signalling these distinctions are traditionally divided into three major inflectional paradigms or 'declensions', one for each gender. The declensions often have more than one ending for a particular case, so further subdivisions are necessary within each major paradigm. Moreover,

[2] Actually, *-n* does have a syllabic allomorph, namely *-en*; but it is evident from the list of non-standard compounds provided by Clahsen *et al.* that their subjects never deleted the latter in compounds.

[3] An interesting attempt at such an explanation is offered by Wegener (1999).

TABLE 4. *Genitive inflections*

Endings	Masculine	Feminine	Neuter	Adjectival
Singular	*-a*, *-u*	*-i*/*-y*	*-a*	*-ego*, *-ej*
Plural	*-ów*, *-y*/*-i*	*Ø*, *-y*/*-i*	*Ø*, *-y*/*-i*	*-ich*/*-ych*

the subdivisions required for one case often cross-cut those required for other cases, so the total number of distinct paradigms is quite large (cf. Orzechowska, 1998).

As the traditional names of the declensions suggest, the most important factor determining the choice of ending is gender, though various considerations come into play when a declension has more than one inflection signalling a particular grammatical distinction. Gender can usually be predicted from the nominative ending. The vast majority of feminine nouns end in *-a* or *-i* in the nominative, though there are a few which end in a palatalized consonant. Masculine nouns nearly always end in a consonant; and neuter nouns usually end in *-o* or *-ę*.

The endings for the genitive inflection are listed in Table 4. The feminine singular ending *-y* and its variant *-i* are used with most feminine nouns and with virile (masculine-human) nouns ending in *-a* or *-o*. Animate masculine nouns and most neuter nouns take *-a*. Inanimate masculine nouns take either *-a* or *-u*, depending on the noun. Which of these two endings a particular noun takes is largely arbitrary, although there are some broad regularities. For example, most nouns which designate tools and body parts, the names of the months, and native place names require the *-a* ending. In contrast, abstract nouns, collective nouns, and mass nouns usually take *-u*, as do most borrowings and most foreign place names. Certain derivational affixes favour one or the other inflection: for example, nouns that end in *-ak* and *-nik* and most diminutives ending in *-ek*, *-ik*, and *-yk* take *-a* (but nouns ending in *-unek* take *-u*). Phonological criteria also play a role (nouns that end in a palatalised consonant nearly always take *-a*). However, there are many exceptions to these tendencies, and some of them are contradictory (see Kottum, 1981).

In the plural, there are also three endings. The regular endings are *-ów* (masculine nouns) and *-Ø* (feminine and neuter nouns), but some nouns of all three genders require *-i*/*-y*, and neuter nouns ending in *-um* take *-ów*. The subclasses defined by the ending they take in the genitive plural cross-cut those found in the singular: for example, masculine nouns that take *-a* in the singular can take either *-ów* or *-i*/*-y* in the plural; neuter nouns pattern with masculines in the singular and with feminine nouns in the plural, and so on. The criteria determining membership in these subclasses are complex, but,

in contrast to those determining the choice of *-a* or *-u* in the masculine declension in the singular, they are describable in general terms (see Orzechowska, 1998).

Further complications are due to the fact that some nouns do not decline at all, while others require endings normally reserved for adjectives. The former category includes borrowings such as *guru* or *whisky*, which do not sound like Polish words, and hence cannot be assimilated into any of the native categories,[4] and many borrowings which could in principle be assimilated but are not (e.g. *jam session*, *rodeo*, *morale*). Nouns that take adjectival endings include deadjectival nouns derived by conversion (e.g. *uczony* 'learned' or 'scholar', *służąca* 'serving (fem.)' or 'maid'), foreign names ending in *-e* (e.g. *Goethe*, *Rilke*), and most native surnames (the *-ski* and *-cki* endings so characteristic of Polish surnames are adjectival suffixes).

Finally, many nouns require various stem changes in some forms. For example, the genitive singular of *cukier* 'sugar' is *cukr-u* (/e/ → Ø); *orzeł* 'eagle' changes to *orł-a* (/ʒ/ → /r/, /e/ → Ø), and *mąż* 'husband' changes to *męż-a* (/õ/ → /ẽ/). Although these alternations are restricted to narrowly defined groups of stems, some of them are quite systematic, while others are truly irregular.

IS THERE A DEFAULT ENDING FOR THE GENITIVE SINGULAR?

We saw in the preceding section that none of the three endings signalling the genitive singular applies to 'nouns in general': *-y* is restricted to feminine nouns and masculine nouns ending in *-a*, *-a* to neuter and masculine nouns, and *-u* to masculine nouns only. This is not a problem for the dual mechanism theory, as rules can be formulated as applying to subcategories of nouns: for example, the *-y* suffixation rule would apply to nouns carrying the feature [FEMININE], and another rule would add *-a* to nouns marked [NEUTER].

However, there remains a problem with the masculine declension, which, as we have seen, has two endings: *-a* and *-u*. Both of these apply to an open-ended class of nouns and both are fairly unrestricted phonologically. Since *-a* is by a large margin the more frequent ending (accounting for 70–80% of masculine types as well as tokens), it is usually considered the 'regular' ending. But according to the dual mechanism theory, regularity in the psychological sense (i.e. default status) is independent of frequency, so in order to determine which, if any, of the two endings is regular, we must consider other criteria.

Marcus *et al.* (1995), as explained earlier, compiled a list of circumstances in which the regular inflection must be used. Since the entire list is assumed

[4] No native nouns end in *-u* in the nominative singular, and the few that end in *-y* are transparently deadjectival.

TABLE 5. *Distribution of* -a *and* -u *endings on masculine nouns in contexts calling for a default inflection*

Circumstance	Required ending	Examples (all forms given in the genitive)
1. Novel words	*-a* *-u*	*frusta*, *bukala*, *milaja* *frustu*, *bukalu*
2. Low-frequency words	*-a* *-u*	*szańca* 'bulwark', *tygla* 'crucible' *częstokołu* 'palisade', *pucharu* 'goblet'
3. Unusual-sounding words	uninflected	*swahili*, *attaché*, *pony*, *Delacroix*
4. Homophones	*-a* *-u*	*tłoka* 'piston', *skręta* 'fag' *tłoku* 'crowd', *skrętu* 'turn'
5. Rhymes	*-a* *-u*	*traktora* 'tractor', *robaka* 'worm' *motoru* 'motorcycle', *baku* 'petrol tank'
6. Onomatopoeia	N/A (renditions of sounds are neuter)	N/A
7. Words mentioned rather than used	Form in which the word was originally used or usual ending	*Nie mogę znaleźć tego drugiego "autorowi"/ "autora"/"rasizmu".* 'I can't find the second (occurrence of the word) 'author (DAT)'/'author'/'racism'.
8. Surnames	*-ego* *-a*	*Bogusławskiego*, *Saloniego*, *Chomsky'ego* *Chopina*, *Mickiewicza*, *Darwina*
9. Borrowings	uninflected *-u* sometimes *-a*	*guru*, *boa*, *kamikadze*, *dingo* *fonemu*, *pubu*, *Wehrmachtu* *drinka*, *jeepa*, *pikadora*
10. Truncations	*-a* sometimes *-u*	*merca* 'Mercedes', *speca* 'specialist' *samu* (from *sklep samoobslugowy*) 'supermarket', *haszu* 'hashish'
11. Acronyms	uninflected *-u*	*PCK* (*Polski Czerwony Krzyż*, 'Polish Red Cross') *PAN-u* (*Polska Akademia Nauk*, 'Polish Academy of Sciences')

12. Derivation from a different category:		
(a) affixation	usually *-a*	*zszywacza* 'stapler' (from *zszywać* 'to stitch together'), *lizaka* 'lollipop' (from *lizać* 'to lick')
(b) backforation	usually *-u*	*zbioru* 'collection' (from *zbierać* 'to collect'), *wykładu* 'lecture' (from *wykładać* 'to lecture')
(c) nominalized adjectives	*-ego*	*chorego* 'ill (person)', *uczonego* 'scholar' or 'learned'
13. Derivation via a different category	*-a*	*utleniacza* 'oxidant' (from *tlen* 'oxygen' via *utleniać* 'oxidize')
	-u	*dodruku* 'additional print run' (from *dodrukować* 'print additional copies' from *drukować* 'to print' from *druk* 'print (n.)')
14. Derivation via name	*-a*	(*ekranizacja*) *Hamleta/Króla Lira*
	-u	*Tajfunu/Wiśniowego Sadu* '(film version of) *Hamlet/King Lear/Typhoon/The Cherry Orchard*'
15/16. Bahuvrihi compounds and nominalized phrases	*-a*	*łamistrajka* 'scab' (lit. 'break-strike'; cf. *strajku* 'strike'), *kątomierza* 'protractor' (lit. 'angle-measure')
	–u	*trójzębu* 'trident' ('three-tooth', cf. *zęba* 'tooth'), *długopisu* 'ball-point pen' (lit. 'long-write')

to fall out automatically from the way the language faculty is organized, it should apply to default inflections in all languages, not just English. The obvious test, then, is to see which of the two endings is required in each of these special situations.

Table 5 shows which ending is used with masculine nouns in the sixteen circumstances which require the default inflection.[5] (Items 15 and 16 are treated as a single category as they are difficult to distinguish in Polish.) As we can see from the table, most of these tolerate both endings; which ending is actually used depends on the lexical properties of the noun. In three circumstances, the noun either can or must be left uninflected; and in two cases, the adjectival ending is called for.

It is clear, then, that there is no single ending applicable in all the circumstances which, according to the dual mechanism theory, call for the regular inflection. This finding is awkward for the theory, but it is certainly not a fatal blow. Although Marcus, Pinker, and their co-authors have argued that these circumstances are associated with the default inflection in every language, at no point did they stipulate that every inflectional category in every language must have a default. It is possible that the Polish masculine declension is exceptional in that it does not have a regular ending for the genitive singular. Since systems as irregular as the Polish genitive singular are rare, this does not seem an unreasonable conclusion.

But the very fact that the genitive singular of masculine nouns lacks a default ending makes it a perfect test case for the dual mechanism theory. The Polish genitive singular might be an oddball among inflectional systems, but Polish speakers are presumably equipped with the same language-processing mechanisms as speakers of other languages. According to the dual mechanism theory, children are programmed to look for default inflections, and are able to identify the default by checking which affix applies in one, perhaps two, of the circumstances listed in Tables 1 and 5 (cf. Marcus *et al.*, 1995: 245). Thus, the theory predicts that Polish children will be prone to misconstruing one of the affixes as the default. (This could be *-a*, *-u*, *-ego*, or *-Ø*, depending on which circumstances actually serve as default

[5] Table 5 lists the various categories of words which, according to the dual mechanism theory, require the regular inflection – i.e. the first sixteen of the original list of twenty-one circumstances compiled by Marcus *et al.* (1995). As noted earlier, the remaining five circumstances involve overgeneralization errors produced by various categories of language users. At the moment, no data is available on overgeneralization of the genitive singular by four of the five populations (Alzheimer's patients, anomic aphasics, Williams syndrome sufferers, and normal adults). Overgeneralization errors in children's speech will be discussed in the following section.

The information given in the table is based on the author's intuitions backed up by lexicographic materials (Szymczak, 1983) and the standard university grammar (Grzegorczykowa, Laskowski & Wróbel,1998), or, in the case of novel words, an informal survey of 18 native speakers.

identifiers – Marcus *et al.*, 1995 do not specify this – and the specific words that the child encounters in these contexts.) In a system as complex and irregular as the Polish genitive singular, this would lead to high overgeneralization rates. Thus, the dual mechanism theory makes two predictions about the acquisition of the genitive singular: children should make relatively frequent overgeneralization errors, and they should consistently overgeneralize one of the endings.

We will now examine some child language data to determine if these predictions are borne out.

LEARNING A SYSTEM WITHOUT A DEFAULT

The data

The corpus of data analysed here contains the spontaneous speech of three of the 'Kraków children': Basia, Inka, and Jaś (Table 6). It was collected in the

TABLE 6. *The data*

Child	Age	Corpus size (child words)
Basia	1;5–4;11	29255
Inka	1;4–4;11	84431
Jaś	1;6–4;11	85856

1950s by the children's mothers as part of a project co-ordinated by Stefan Szuman, and subsequently digitalized and tagged by Magdalena Smoczyńska (see Smoczyńska, 1998). The children were observed at frequent intervals, often daily, over a period of years, and their speech was sampled in a variety of situations to obtain a representative record of their linguistic abilities. The three children chosen for this study were those whose records spanned the entire period from the emergence of the genitive to age 4;11.

Overall course of development

At first glance, the development of the genitive singular inflection in the three children follows a familiar developmental pattern: a slow beginning, followed by a period of inconsistent usage, followed by a sudden improvement in performance, presumably corresponding to rule acquisition. What is surprising, in view of the complexity of the Polish case system, is the fact that the acquisition occurs so early: all three children reach Brown's criterion (90% correct in obligatory contexts) before age 2;0. Even more surprising is the fact acquisition of the genitive singular is virtually errorless. The children sometimes use the citation form (the nominative) rather than the genitive, particularly in the earlier transcripts; and there are some instances of the accusative being used after negated verbs, which require genitive objects.

However, when the genitive is supplied, it is nearly always the correct form (see Tables 7 and 8).

TABLE 7. *Correct forms and errors in genitive singular contexts*

Child	Correct	Citation form	OG errors	Other errors	Total
Basia	942	55	5	25	1027
Inka	2719	30	17	44	2810
Jaś	2902	22	60	56	3040
Total	6563	107	82	125	6877

TABLE 8. *Acquisition of the genitive singular : summary*

Child	Emergence	Acquisition	First OG	Mean OG rate	Maximum OG rate
Basia	1;5	1;10	2;6	0·53%	2·56%
Inka	1;4	1;9	1;10	0·62%	1·41%
Jaś	before 1;6	1;11	1;9	2·02%	4·85%

Overgeneralization errors appear late and are rare. This is particularly striking in the case of Basia. The first recorded overgeneralization in her corpus occurred at 2;6, more than a year after the first genitive. Before that point, as far as we can tell from the transcripts, she never used the feminine ending with a masculine or neuter noun or vice versa; she never used the *-a* ending with a masculine noun which required *-u*; she never attempted to add an ending to a noun that does not inflect; and she never failed to make stem alternations when these were required. Basia's transcripts for the entire period studied contain 947 explicitly marked genitives, only 5 of which are overgeneralizations, giving a mean (absolute) overgeneralization rate of 0·53%. Her highest overgeneralization rate in any three-month period is 2·56%.

Given the rarity of overgeneralization errors in Basia's speech, their apparently late appearance could be a sampling artefact. However, we can be reasonably confident that the unrecorded overgeneralizations did not occur much earlier than 2;6. Basia's records for the period up to the first overgeneralization contain 409 genitive tokens of 94 types, all correct. Then, during the three-month period from 2;6 to 2;8, there are 76 correct forms, two overgeneralization errors, and four instances of a singular stem misanalyzed as a plural form. Following this small burst of inflectional errors, there are only three further overgeneralizations, one at 3;8, one at 4;0, and one at 4;5.

Inka began to use the genitive at 1;4, and used it consistently (in over 90% of obligatory contexts) from 1;9. The first recorded overgeneralization

TABLE 9. *Overgeneralization errors in the genitive singular*

Child	*-a*	*-u*	*-y/-i*	Wrong stem	Total
Basia	2	1	0	2	5
Inka	9	2	3	3	17
Jaś	39	10	1	10	60
Total	50	13	4	15	82

TABLE 10. *Across-declension and within-declension errors in the genitive singular*

	Across-declension errors			Within-declension errors	
Child	*-a*	*-u*	*-y/-i*	*-a*	*-u*
Basia	0	1	0	2	0
Inka	0	1	3	9	1
Jaś	10	1	1	29	9
Total	10	3	4	40	10

occurred a month later, at 1;10. Her highest overgeneralization rate in any three-month period (1·41 %) is even lower than Basia's maximum, though her mean rate (0·62 %) is slightly higher. Her overall development is rather similar to Basia's, though not quite so extreme.

The third child, Jaś, was already using the genitive at 1;6, when his records begin, but it is unlikely that he had been doing so for very long, as citation forms in genitive contexts are still quite frequent. He reliably supplied the correct ending when required from 1;11 onwards. His first overgeneralization error occurred very early, at 1;9. He had the highest overgeneralization rate of the three children: mean 2·02 %, with a maximum at 4·85 %.

Error analysis

We now turn to an analysis of the children's overgeneralizations errors. Overgeneralization errors arise when the child either uses an inappropriate ending (e.g. a masculine-declension ending with a feminine noun) or the correct ending with the wrong stem (i.e. when the child fails to make the necessary stem changes). As we can see from Table 9, the children made errors of both kinds, and they overgeneralized all three endings. On the other hand, it is also clear that *-a* was overgeneralized more frequently than the other affixes. Does this mean that the children were treating it as the default? Several lines of evidence seem to argue against such an interpretation.

First, *-a* was not overextended indiscriminately to all nouns: by and large, *-a* overgeneralizations were confined to masculine nouns that require *-u*. We

can see this when we compare the number of within-declension errors (errors involving confusion of the two masculine endings) and across-declension errors (involving the use of an ending characteristic of one declension with a noun belonging to another, e.g. a feminine-declension ending with a masculine noun or vice versa).[6] As we can see from Table 10, almost three quarters of the errors with endings occur within the masculine declension. Across-declension errors are rare, and they are just as likely to involve *-u* or *-y* as *-a*. What is more, they tend to be found only in the earlier transcripts or with exceptional nouns (feminine nouns that sound like masculines or vice versa). It is clear, then, that children learn to restrict endings to nouns of the appropriate gender very early in acquisition. The distribution of *-a* and *-u* within the masculine declension causes more problems because it is largely arbitrary.

Secondly, although all three children overgeneralized *-a* more frequently than *-u*, the differences in OG rates for the two affixes are much smaller than the differences in the frequency of regularization and irregularization errors reported for English-speaking children. Furthermore, the children's preference for *-a* can be confidently attributed to the fact it is the more frequent of the two endings. This is evident from the fact that relative overgeneralization rates for the two affixes are very similar (cf. Table 11;[7] the

TABLE 11. *Absolute and relative overgeneralization rates in the genitive singular*

	Tokens correct			Overgeneralizations			OG rate (%)			ROG rate (%)		
Child	*-a*	*-u*	*-y/-i*	*-a*	*-u*	*-y/-i*	*-a*	*-u*	*-y/-i*	*-a*	*-u*	*-y/-i*
Basia	387	118	422	2	1	0	0·37	0·12	0·00	0·51	0·84	0·00
Inka	1050	424	1187	9	2	3	0·56	0·09	0·20	0·85	0·47	0·25
Jaś	1137	429	1264	39	10	1	2·25	0·41	0·06	3·32	2·28	0·08
Total	2574	971	2873	50	13	4	1·28	0·24	0·11	1·91	1·32	0·14

ROG formula, as explained in the section on German plurals, factors out the effect of frequency). In absolute terms, *-a* overgeneralizations outnumber *-u* errors by approximately 3·4:1. This figure is almost exactly the same as the ratio of *-a* to *-u* genitives in the children's correct productions (3·3:1).

Finally, it is worth pointing out that both OG and ROG rates for the one inflection which is undeniably regular, the feminine *-y*, are much lower than those for the other two endings. This is not surprising, since, unlike the other

[6] Errors involving confusion of the two variants of the feminine endings (*-i* and *-y*) are extremely rare and difficult to distinguish in a principled way from non-standard pronunciations, so they will not be considered here.

[7] The data in Table 11 do not include nouns which take adjectival endings and nouns which do not decline. Stem errors have also been excluded.

two endings, it has a clear domain of application; but it is difficult to reconcile with a central tenet of the dual mechanism theory – namely, the claim that regular inflections are much more prone to overgeneralization than irregular patterns.

There is no evidence, then, of a dissociation between regular and irregular endings. Polish children acquire three productive patterns, each restricted to a particular subset of nouns; and they occasionally overgeneralize all three. There is little to suggest that different mental mechanisms are involved.

COMPARISON WITH THE ACQUISITION OF THE ENGLISH PAST TENSE

The course of development that we observe in the three Polish children (and especially in Basia) is very different from the familiar pattern of morphological development in English-speaking children. Perhaps the most obvious difference is in age of acquisition: like other children acquiring heavily inflected languages, Polish children learn the basic inflectional patterns much earlier than English-speaking children. The children studied here began to use the genitive singular between 1;4 and 1;6, and used it consistently (i.e. in over 90% of obligatory contexts) about three months later. English-speaking children reliably mark tense from about three (Brown 1973; Kuczaj, 1977); the possessive inflection is acquired at about the same time (see Brown, 1973; de Villiers & de Villiers, 1973). This difference is probably due to the fact that inflections in Polish carry a greater functional load than they do in English (cf. Smoczyńska, 1985).

Secondly, Polish children appear to be more cautious. Most English-speaking children begin to overregularize at a time when they are not yet consistently supplying the regular inflection in all obligatory contexts. For example, on examining Cazden's data for the Harvard children, we find that in the month in which the first regularization error occurred, Adam provided the correct regular form in 45% of the obligatory contexts, Eve in 20% and Sarah in 71%. Moreover, in all three cases, several months elapsed before the children consistently supplied the regular inflection in obligatory contexts (Cazden's data, cited by Marcus *et al.*, 1992, Appendices A1–A3). There are also several examples in the literature of children whose very first recorded *-ed* past tense form was a regularization (Ervin, 1964: 178; Marcus *et al.*, 1992: 105). In contrast, two of the Polish children, Basia and Inka, were already reliably marking the genitive at the time the first overgeneralization errors were recorded, while Jaś marked the genitive 82% of the time in the month in which the first overgeneralization errors occurred.

Another important difference is that Polish children overgeneralize all three endings, not just a single 'default' ending. It is true that one of the endings (*-a*) is overgeneralized more frequently than the other two, but there is little evidence to suggest that it is used as a default. It is no more likely to be extended to declension-inappropriate contexts than the other two endings,

and the differences in the frequency with which each of the three endings is overgeneralized are small in comparison to the dramatic contrast between the frequency of regularization and irregularization errors in English.

Finally, overgeneralization rates for the Polish children are very low. OG rates for the feminine ending *-y* range from 0 to 0·2%, with a mean rate at 0·11%. Even *-a*, the most troublesome ending, was overgeneralized in only 1·31% of the opportunities. As explained earlier, estimates of the frequency of overgeneralization of the regular ending by English speaking children vary considerably from study to study, but it seems that somewhere between 4% and 10% is a reasonable estimate.

The lower rates of overgeneralization in the Polish-speaking children are puzzling, since the Polish system is clearly much more complex and more irregular. Since overgeneralization rates vary considerably between children, it is tempting to dismiss this finding as a mere coincidence: the three children might have been particularly cautious learners with unusually low overgeneralization rates. While this possibility cannot be ruled out, it could only provide a partial explanation for this unexpected finding, as all three children had very high overgeneralization rates in the genitive plural.

THE GENITIVE PLURAL

In fact, the development of genitive plural inflections follows the 'English' pattern (see Table 12). The genitive plural is usually acquired later, which is

TABLE 12. *Acquisition of the genitive plural: summary*

Child	Emergence	Acquisition	First OG	Mean OG rate	Maximum OG rate
Basia	1;11	after 5	2;10	18·82%	33·33%
Inka	1;7	2;6	2;9	4·20%	10·00%
Jaś	1;9	4;0	1;11	11·04%	21·74%

not surprising, since it is more complex semantically and less frequent than the singular. Overgeneralization errors are much more frequent (Figure 1), and they appear either at the same time or before the child reliably supplies the inflection. Most importantly, one ending, *-ów*, is consistently overgeneralized: *-ów* overgeneralizations accounted for all 32 of Basia's overgeneralization errors in the plural, 11 out of the 12 produced by Inka, and 55 out of the 67 that Jaś made (see Table 13). Moreover, in contrast to the singular endings, *-ów* is frequently used in declension-inappropriate contexts. In short, the children appear to treat it as a 'default' ending.

So we have a different developmental pattern for the genitive in the singular and the plural IN THE SAME CHILDREN. Clearly, this effect cannot be argued away by appealing to individual acquisition strategies: the source of

TABLE 13. *Absolute and relative overgeneralization rates in the genitive plural*

	Tokens correct			Overgeneralizations			OG rate (%)			ROG rate (%)		
Child	*-ów*	*-Ø*	*-i/-y*	*-ów*	*-Ø*	*-i/-y*	*-ów*	*-Ø*	*-i/-y*	*-ów*	*-Ø*	*-i/-y*
Basia	57	45	32	32	0	0	29·36	0·00	0·00	35·96	0·00	0·00
Inka	83	98	81	11	0	1	5·79	0·00	0·55	11·70	0·00	1·22
Jaś	238	204	73	55	10	2	16·57	3·12	0·45	18·77	4·67	2·67
Total	378	347	186	98	10	3	15·53	1·74	0·41	20·59	2·80	1·59

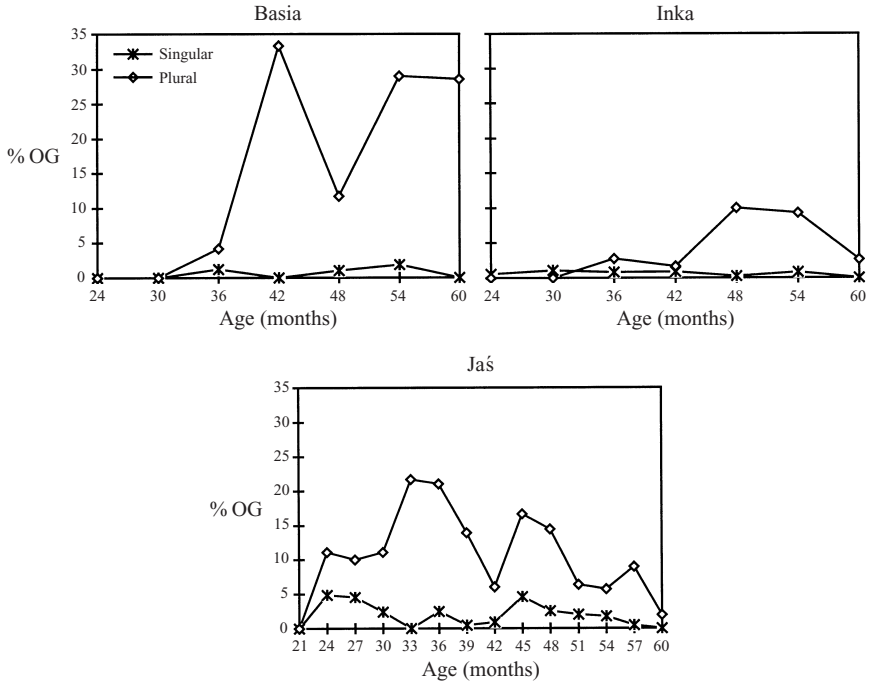

Fig. 1. Overgeneralization rates in the genitive singular and plural.

the differences must reside in the linguistic system. We know from earlier studies that the overgeneralization of the genitive plural *-ów* is the only inflectional error that is both common and persistent, and it is also the only case ending which is frequently used in declension-inappropriate contexts (see Smoczyńska, 1985). This suggests that the source of the difficulty lies in some idiosyncratic property of the genitive plural inflection. What is it, then, that sets the genitive plural apart from the rest of the case-marking system?

Inflected forms normally consist of a stem and an ending. In the genitive

plural, there are three main endings: *-ów*, *-i*/*-y*, and -∅. The *-i*/*-y* ending is fairly rare, phonologically restricted, and has low cue validity (it also marks the nominative, genitive, dative, locative, and vocative singular and the nominative, accusative, and vocative plural of some nouns), so in effect the child must choose between *-ów* and -∅. The zero forms are unusual in that they do not contain overt case markers, although the absence of an ending often triggers stem alternations which serve as secondary markers of the genitive. For example, the genitive plural of *kobieta* 'woman' is *kobiet* (zero ending, no stem changes); but in *nóg* (gen. pl. of *noga* 'leg') the stem vowel /o/ changes to /u/, in *rąk* (gen. pl. of *ręka* 'hand') the stem vowel /ẽ/ changes to /õ/, and in *piłek* (gen. pl. of *piłka* 'ball') a vowel is inserted to break up a word-final consonant cluster. The stem alternations are rule-governed – that is to say, they are obligatory in certain contexts – but they are restricted to groups of stems with fairly specific phonological characteristics. Furthermore, different families of stems undergo different alternations.

Thus, the zero-marked forms are unsegmentable and often shorter than the nominative, which does contain an affix. In effect, the genitive marker is buried inside the stem. In contrast, the *-ów* suffixed forms contain an explicit (and very prominent) marker of the genitive. Not surprisingly, the high error rates in the genitive plural are due almost entirely to overgeneralization of the more explicit mechanism.

The parallels with the English system of marking past tense are striking.[8] Regular past tense forms in English, like *-ów* genitives, consist of a stem and an affix. The past tense of irregular verbs, in contrast, is signalled by stem changes – several different stem changes. The fact that Polish children are prone to overgeneralizing an affix when it competes with stem changes but not when it competes with other affixes suggests that affixation is inherently more generalizable than stem changes, confirming the suspicion expressed earlier that the sharp contrast between regular and irregular verbs in English may be partly attributable to the fact that they rely on different morphological mechanisms to mark tense.

CONCLUSION

The Polish system of marking the genitive singular, as we have seen, is complex and highly irregular. In spite of this, it is acquired early and almost

[8] On the other hand, it must be emphasized that the parallelism between the English past-tense and Polish genitive plural inflections is only partial. In addition to the obvious difference – the fact that one signals tense and the other case relations – there are also differences in how these distinctions are marked. Polish children's propensity to overgeneralize *-ów* is without doubt amplified by its syllabic status, as well as the non-canonical nature of the competing zero-inflected forms. (All other inflected forms are a syllable longer, and consequently zero-marked genitives are perceived as 'too short': they appear to be missing a syllable – cf. Smoczyńska, 1986.) The English regular past tense ending, usually realised by a single alveolar segment, is much less salient.

without error. The three children in this study appeared to have little trouble in restricting each ending to a particular subset of nouns, even when this subset was essentially arbitrary, and there is no evidence that they treated any one ending as the default. Clearly, we cannot draw any far-reaching conclusions on the basis of a small study involving only three children; but the findings described here do raise certain problems for the dual mechanism theory.

To accommodate the data on the genitive singular in the theory, we must assume that children acquire this inflection using associative memory alone. But according to proponents of the dual mechanism theory, children actively look for the default, and are able to identify it by checking which inflection is used in some tell-tale context (see Marcus *et al.*, 1992, 1995; Clahsen, 1999). If this is so, Polish-speaking children could easily conclude that whichever ending they happened to hear in the relevant context was the category default; and it is not clear how they would recover from such an error. Furthermore, the ease and speed of acquisition of the genitive singular seems to make the default mechanism redundant: if children can acquire such a complex system so quickly without the benefit of the default mechanism, one wonders whether they really need it to learn a relatively simple system such as the English past tense.

The acquisition of the genitive plural, in contrast, resembles the familiar developmental pattern found in English-speaking children learning to form the past tense. Overgeneralization errors in the plural are comparatively frequent and one ending is applied indiscriminately to nouns of all genders. We saw that this pattern of development is untypical for Polish, and that it can be attributed to the idiosyncratic properties of the genitive plural inflection.

This supports the suspicion expressed earlier that it is the special properties of the English system of marking tense that are responsible for the sharp dissociations between the regular and the irregular inflections that proponents of the dual mechanism theory so painstakingly documented. These properties – the asymmetry in frequency between the regulars and the irregulars, the fact that they rely on different morphological mechanisms, that irregulars form phonologically similar clusters, and that there is only one truly productive inflection – exaggerate the differences between regular and irregular forms, and their combined effect is to make what may be no more that a quantitative difference appear qualitative.

That each of these properties individually contributes to generalizability can be demonstrated by comparing overgeneralization rates for inflections which differ with respect to a single property only. The role of frequency is revealed by the two masculine singular endings of the Polish genitive, *-a* and *-u*. Both of these use the same morphological mechanism, suffixation, and both are phonologically unrestricted, but *-a*, the more common ending, is

overgeneralized more frequently than its rarer counterpart, *-u*. In fact, as pointed out earlier, the difference in the frequency of overgeneralization of the two endings is almost exactly the same as the difference in the frequency of *-a* and *-u* in the children's correct productions. The German findings summarized in section 3 also confirm this effect: German-speaking children most frequently overgeneralize the plural ending with the highest type frequency, namely *-n*.

The children's errors in the genitive plural highlight the effect of transparency. With some nouns, the genitive plural is marked by an overt ending (*-ów* or *-i*/*-y*); with others, the same grammatical function is conveyed by the absence of an affix, which often triggers stem alternations. The zero ending and *-ów* have similar frequencies, yet children show a very strong preference for the more transparent forms with an explicit suffix. Research on the acquisition of other morphological systems provides additional evidence that transparency affects generalizability. Orsolini, Fanari & Bowles (1998) report that Italian children overgeneralize irregular verbal endings much more frequently than irregular stem changes; and German children seem to prefer affixes to umlauts as plural markers (Köpcke, 1998).

The effect of applicability is best illustrated by the German plural affixes. Four of the affixes (*-en*, *-er*, *-e*, and *-Ø*) are associated primarily with subsets of stems sharing certain phonological and grammatical features, while the fifth, *-s*, is much less discriminating. German children overgeneralize the *-s* ending much more readily than we would expect on the basis of its frequency alone, showing that wide applicability also contributes to generalizability. Conversely, Polish children very rarely overgeneralize the genitive singular *-y* ending. This is without doubt attributable to the fact that it has a very clear domain of applicability: feminine nouns, the vast majority of which end in *-a*, and masculine nouns ending in *-a*.

The last of the factors responsible for the sharp contrast between regulars and irregulars in English is the fact that there is only one truly productive ending, *-ed*. The various strong verb patterns are, to all practical intents and purposes, restricted to groups of specific WORDS, so when speakers are presented with a word that does not belong to any of these groups, they have no choice but to use the regular inflection. On the other hand, when the language offers viable alternatives, speakers make use of them: Polish children learning the genitive and German children learning the plural generalize, and overgeneralize, all the patterns they find in the language, not just a single default.

The English system, with its single 'default' ending, then, is a special case. This does not mean that it is exceptional: the particular combination of properties that we find in the English tense marking system is not uncommon. In languages that use both stem changes and affixation, the latter is usually more productive since, for historical reasons, stem changes tend to be

phonologically restricted (cf. Bybee, 1995); and irregular forms typically have high token frequencies, simply because those that do not are prone to regularization. The English past tense is a special case in the sense that it is just one of the many types of morphological systems that the child might encounter in the input. In the languages of the world, we find a whole spectrum of possibilities. Some, like the English past tense, use just a single regular ending; others, like the Polish genitive singular of masculine nouns, use several endings with largely arbitrary distributions; and still others occupy various intermediate points on this continuum.

The empirical evidence for the dual mechanism theory comes in large part from studies of the acquisition and processing of the English past tense and plural inflections. However, an adequate psycholinguistic theory must be able to account for data from any language, not just English. The data on the acquisition of the Polish genitive singular and the German plural reviewed here, as well as Orsolini's work on the acquisition of Italian verbal morphology (Orsolini *et al.*, 1998), show little evidence of a categorical contrast between a fully productive default inflection and highly restricted irregular patterns. Clearly, more research is necessary before we reach any firm conclusions, but the results from languages with more elaborate inflectional systems suggest that the dual mechanism theory may be due for a reassessment.

REFERENCES

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

Bittner, D. & Köpcke, K.-M. (2000). Acquisition of the German plural markings: a case study in natural and cognitive morphology. In C. Schaner-Wolles, J. R. Rennison & F. Neubarth (eds), *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler Presented on the Occasion of his 60th Birthday*. Torino: Rosenberg & Sellier.

Brown, R. (1973). *A First Language. The Early Stages*. Cambridge, MA: Harvard University Press.

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., Rothweiler, M., Woest, A. & Marcus, G. F. (1992). Regular and irregular inflection in the acquisition of German noun plurals. *Cognition* **45**, 225–55.

de Villiers, J. & de Villiers, P. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research* **2**, 331–41.

Ervin, S. M. (1964). Imitation and structural change in children's language. In E. Lenneberg (eds), *New Directions in the Study of Language*. Cambridge, MA: MIT Press.

Gordon, P. (1985). Level-ordering in lexical development. *Cognition* **21**, 73–93.

Grzegorczykowa, R., Laskowski, R. & Wróbel, H. (1998). *Gramatyka współczesnego języka polskiego. Morfologia*. Warszawa: PWN.

Hare, M., Elman, J. L. & Daugherty, K. G. (1995). Default generalisation in connectionist networks. *Language and Cognitive Processes* **10**, 601–30.

Jaeger, J. J., Lockwood, A. H., Kemmerer, D. L., Van Valin, R. D. & Murphy, B. W. (1996). Positron emission tomographic study of regular and irregular verb morphology in English. *Language* **72**, 451–97.

Kiparsky, P. (1985). Some consequences of lexical phonology. *Phonology Yearbook* **2**, 85–138.

Köpcke, K.-M. (1988). Schemas in German plural formation. *Lingua* **74**, 303–35.
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.
Kuczaj, S. A. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior* **16**, 589–600.
Marchman, V. A. (1997). Children's productivity in the English past tense: the role of frequency, phonology, and neighborhood structure. *Cognitive Science* **21 (3)**, 28–0.
Marcus, G. F. (1995). Children's overregularization of English plurals – a quantitative analysis. *Journal of Child Language* **22**, 447–59.
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**, 1–180.
Marcus, G. F., Brinkman, U., Clahsen, H., Wiese, R. & Pinker, S. (1995). German inflection: the exception that proves the rule. *Cognitive Psychology* **29**, 189–256.
Mills, A. (1985). The acquisition of German. In D. I. Slobin (ed.), *The cross linguistic study of language acquisition. Volume 1 : The data*. Hillsdale, NJ: Erlbaum.
Orsolini, M., Fanari, R. & Bowles, H. (1998). Acquiring regular and irregular inflection in a language with verb classes. *Language and Cognitive Processes* **13**, 425–64.
Orzechowska, A. (1998). Rzeczownik. In R. Grzegorczykowa, R. Laskowski & H. Wróbel (eds), *Gramatyka współczesnego języka polskiego : Morfologia*. Volume 1. Warszawa: PWN.
Park, T. Z. (1978). Plurals in child speech. *Journal of Child Language* **5**, 237–50.
Pinker, S. (1997). Words and rules in the human brain. *Nature* **387**, 547–8.
Pinker, S. (1998). Words and rules. *Lingua* **106**, 219–42.
Pinker, S. (1999). *Words and rules. The ingredients of language*. London: Weidenfeld and Nicolson.
Pinker, S. & Prince, A. (1988). On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* **28**, 73–193.
Pinker, S. & Prince, A. (1992). Regular and irregular morphology and the psychological status of rules of grammar. *Berkeley Linguistics Society* **17**, 230–51.
Plunkett, K. & Marchman, V. (1993). From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition* **48**, 21–69.
Pulvermüller, F. (1998). On the matter of rules – past-tense formation and its significance for cognitive neuroscience. *Network-computation in neural systems* **9**, R1–R52.
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.
Seidenberg, M. S. & Hoeffner, J. H. (1998). Evaluating behavioral and neuroimaging data on past tense processing. *Language* **74**, 104–22.
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. (1986). Analysis of children's errors: some methodological issues. In I. Kurcz, G. W. Shugar & J. H. Danks (eds), *Knowledge and language*. Amsterdam: North Holland.
Smoczyńska, M. (1998). Krakowskie dane językowe dzieci w systemie CHILDES. In M. Smoczyńska (ed.), *Studia z psychologii rozwojowej i psycholingwistyki*. Kraków: Universitas.
Szymczak, M. (1983). *Słownik języka polskiego*. Warszawa: PWN.
Wegener, H. (1994). Variation in the acquisition of German plural morphology by second language learners. In R. Tracy & E. Lattey (eds), *How tolerant is universal grammar? Essays on language variability and language variation*. Tübingen: Niemeyer.
Wegener, H. (1999). Die Pluralbildung im Deutschen – ein Versuch im Rahmen der Optimalitätstheorie. *Linguistik online* **4(3)**, http://viadrina.euv-frankfurt-o.de/~wjournal/3_99/wegener.html.
Xu, F. & Pinker, S. (1995). Weird past tense forms. *Journal of Child Language* **22**, 531–56.