Early noun plurals in German: regularity, productivity or default?*

SABINE LAHA
Austrian Academy of Sciences

DORIT RAVID
Tel Aviv University, Israel

KATHARINA KORECKY-KRÖLL
Austrian Academy of Sciences

GREGOR LAHA
Boku Vienna

AND

WOLFGANG U. DRESSLER
Vienna University

(Received 7 March 2003. Revised 13 June 2005)

ABSTRACT

The acquisition of German plurals has been the focus of controversy in the last decade. In this paper we claim that degree of productivity (i.e. the capacity of nouns to form potential plurals) plays a key role in determining pace of acquisition. A plural elicitation task was administered to 84 Viennese German-speaking children aged 2;6 to 6;0. Analyses of correct responses showed that the highest scores were obtained with -e plurals, followed by the plural markers -e + U, -er

[*] We are grateful to the children who took part in our study, to their families, and to the staff of the kindergartens where the study was conducted. We would also like to thank two anonymous JCL reviewers for comments and helpful suggestions. An earlier version of this paper was presented at the IX International Congress for the Study of Child Language, Madison, July 16–21, 2002. This study was supported by the Austrian Academy of Sciences and by the Austrian Fonds zur Förderung der wissenschaftlichen Forschung FWF (P 17276-G03). Address for correspondence: Sabine Laaha, Department of Linguistics and Communication Research, Austrian Academy of Sciences; c/o Department of Linguistics, Berggasse 11/2/3, 1090 Vienna, Austria. e-mail: sabine.laaha@oeaw.ac.at

271
+U, -s and -(e)n. The lowest score was observed for pure Umlaut (U) plurals. Analyses suggested an impact of productivity on the number of correct scores: fully productive and productive plural patterns obtained higher correct scores than weakly productive and non-productive ones. The results of the study support our productivity scale and are compatible both with single-route models and with a race-model variant of the dual-route view.

INTRODUCTION

The acquisition of inflectional morphology has been the subject of numerous studies since it was first demonstrated that English-speaking children could produce nonce past tense and plural forms by rule. Inflectional morphology is semantically regular and predictable, obligatory and generally applicable, and is therefore produced early on in child language (Brown, 1973). However, not all inflected forms are formally regular. English noun plurals and past-tense verbs have been of particular interest to child language researchers because they are expressed by two different devices: one is regular grammatical inflection, with general application to the whole relevant category – for example, plural keys, past-tense jumped. Another is irregular inflection, which applies to a small number of words within each category, with different degrees of irregularity. For example, the plural forms children of child, men of man are completely irregular and unpredictable, while tooth/teeth, goose/geese, foot/feet share some similarities and cannot be said to be completely irregular. In the same way, past tense forms do/did and bring/brought are complete irregulars, while keep/kept, blow/blew, and sing/sang each represents a small class of sub-regularities. Studies of the route that English-speaking children take in acquiring such systems typically show a U-shaped curve: the child first produces both inflected and uninflected forms (e.g. break/broke), then starts producing both correctly inflected forms as well as overgeneralization in both directions (break/breaked and drag/drug) (Bybee & Slobin, 1982; Marcus, Pinker, Ullman, Hollander, Rosen & Xu, 1992; Marcus, Brinkmann, Clahsen, Wiese & Pinker, 1995) and finally masters both ‘regular’ and ‘irregular’ inflection successfully (Plunkett & Marchman, 1991). However, inflectional systems such as the German noun plural marking are fraught with complexities which prevent their early mastery (Clahsen, Rothweiler, Woest & Marcus, 1992; Köpcke, 1998; Wegener, 1999).

The current study examines the acquisition of noun plurals in German-speaking preschoolers living in Vienna, Austria. German pluralization involves four overt plural suffixes and a zero plural form, in addition to stem alternations (umlaut of the last full vowel or diphthong). Since the frequent
plural markers have a restricted domain, and the least restricted (or ‘default’) marker (-s) is infrequent, the opportunity is provided to trace children’s strategies in learning this complex system in view of current theories of morphological acquisition.

Models of processing and acquiring inflectional morphology

Two current models of morphological representation and processing provide different accounts for the acquisition of inflectional systems. According to the words and rules model (Pinker & Ullman, 2002a,b) the inflectional divide in English represents an epiphenomenon of the design of the human language faculty, where two systems – lexicon and grammar – overtly express inflectional features. This dual-route view (see also Marcus et al., 1992, 1995; Clahsen, 1999; Pinker, 1999, among others) assumes that regular forms are computed in the grammatical portion of the procedural system by productive, combinatorial operations that assemble morphemes and simplex words into complex words and larger syntactic units. Regular forms are thus productively generated by rules which are discrete, categorical and symbolic objects used in a specialized, innate language module. The regular plural is created by a general operation of unification concatenating plural -s with the symbol N and inflecting any word categorized as a noun. An important feature of this view is the dissociation of singular stem and suffix as distinct symbolic variables (Berent, Pinker & Shimron, 2002).

Under this view, irregular forms behave like words in the lexicon, and are acquired and stored like other words with the plural grammatical feature incorporated into their lexical entries. Learning irregular forms is governed by the property of memory which facilitates the acquisition of similar items and superimposes the properties of old items on new ones resembling them. A stored inflected form blocks the application of the rule to that form, but elsewhere the rule applies to any item with the appropriate symbolic category N.

This is not absolute: Pinker & Ullman (2002a) posit a parallel-race model, where regular forms might be stored and accessed depending on specific factors. In the race-model variant of the dual-route view (e.g. Baayen & Schreuder, 1999), it is assumed that the parsing route and the direct route via full forms are both active in parallel. Various factors such as the frequency of an inflected word form, its phonological and semantic transparency, its morphological structure, and lexical neighbourhood determine which of the two routes is more efficient. Based on these assumptions, Pinker & Ullman (2002a: 458) propose that regular forms that resemble irregulars are stored, because the forms must overcome a partial blocking effect exerted by similar irregulars.
Acquiring inflectional morphology under this account would start with memorizing regular and irregular inflected forms which occur in the input. Children would eventually construct or discover the regular rule, which would result in both correct forms as well as some overgeneralizations such as past-tense *holded* or plural *mans*. This is because the symbolic rule is applied across the board in the relevant category, and in some cases blocking of the regular by the irregular form does not occur since young children do not yet have deep memory traces of all irregular forms. With time and exposure to more input, overregularizations disappear.

A second account denies any split in processing mechanisms, and proposes a single-route model to handle both types of inflectional morphology (McClelland & Patterson, 2002). Starting with Rumelhart & McClelland (1986), this model challenges the need for the use of rules (see also Plunkett & Marchman, 1991; Daugherty & Seidenberg, 1994; Bybee, 1995, among others). McClelland & Patterson (2002) propose that both regular and irregular aspects of inflection emerge from a single, integrated mechanism, a parallel distributed processing or connectionist framework in which cognitive processes are viewed as graded, probabilistic, interactive, context-sensitive and domain-general. Under this view, no actual rules operate in the processing of language. Language acquisition and acquisition of other abilities occur via gradual adjustment of the connections among simple processing units, so that what looks like rule-governed performance is in fact approximate descriptions of patterns of language use. The learning mechanism is a pattern-associator network that learns the relationship between phonological forms of the stems and inflected forms of English words, flanked by a fixed encoding network on the input side and a fixed decoding network on the output side. According to McClelland & Patterson, rule application by a dual-route model would be uniform in application to the designated category and independent of contextual factors such as semantics, phonology or frequency. In contrast, a single-route model asserts that acquisition and processing are probabilistic and context-sensitive and thus affected by phonological, semantic and other factors, and that learning is protracted and gradual.

Under this view, both regular and irregular forms would be acquired by the pattern associator in the same way: the network is presented with each pair of singular stem and its plural form and encodes them. Weights between the members of each pair are reviewed and revised with each new presentation—regular and irregular alike—governed by factors such as frequency, similarity, consistency. Later single-route models use the back-propagation algorithm that allows recurrent connections and more layers of hidden units. The learning network improves performance over many learning trials, resulting in a gradual developmental process where overgeneralization is conditioned by linguistic experience coupled with the
similarity of the exemplar being learned to others already stored and by frequency factors.

This paper describes the route taken by young speakers of German learning to pluralize nouns in the perspective of the models described above.

**German noun plurals**

Unlike English inflection, the system of nominal pluralization in German consists of a number of phonologically unrelated plural allomorphs with no clearly dominant form (Köpcke, 1998). German noun plurals are formed by four different suffixes (-s, -(e)n, -e, -er) or by a zero morpheme. Two of these plural markers (-e, zero) may combine with Umlaut (stem vowel change); with -er plurals, the use of Umlaut – if applicable, i.e. to an unumlautable vowel – is obligatory, thus rendering the seven plural markings presented in Table 1.

The assignment of the plural markers in German is largely bound to certain lexical, phonological and (sometimes) semantic characteristics of the nouns (see Köpcke, 1998; Wegener, 1999). The suffixes -(e)n, -e and the zero morpheme are the most frequent plural markers in the adult spoken system, both in type and in token frequency (Goebel & Indefrey, 2000: 180).

One view of German pluralization adopts the dual-route model and assumes a regular/irregular split between -s and all other plural formations (e.g. Marcus et al., 1995; Clahsen, 1999; Pinker, 1999). Marcus et al. (1995) claim that the -s plural, despite its low type and token frequency (1–3%), is the default plural in German: the use of -s is morphophonologically free and it is chosen in default circumstances (p. 229). They present 21 default conditions (p. 240) in which -s plural operates because memory retrieval is blocked, such as with proper names (e.g. Mann ‘man’, Pl. Männ-er, but the

---

**Table 1. The seven plural markers in German**

<table>
<thead>
<tr>
<th>Plural marker</th>
<th>Singular</th>
<th>Plural</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>-s</td>
<td>Auto</td>
<td>Auto-s</td>
<td>cars</td>
</tr>
<tr>
<td>-(e)n</td>
<td>Katze</td>
<td>Katze-n</td>
<td>cats</td>
</tr>
<tr>
<td>-e</td>
<td>Bus</td>
<td>Buss-e</td>
<td>buses</td>
</tr>
<tr>
<td>-e + U</td>
<td>Zug</td>
<td>Züg-e</td>
<td>trains</td>
</tr>
<tr>
<td>zero</td>
<td>Pullover</td>
<td>Pullover</td>
<td>pullovers</td>
</tr>
<tr>
<td>U</td>
<td>Apfel</td>
<td>Äpfel</td>
<td>apples</td>
</tr>
<tr>
<td>-er + U</td>
<td>Haus</td>
<td>Häus-er</td>
<td>houses</td>
</tr>
</tbody>
</table>
brothers Thomas and Heinrich Mann are die Mann-s), derivations from other categories (e.g. nominalized conjunctions die Wenn-s und Aber-s ‘the if-s and but-s’), words with a non-canonical root such as loan words (e.g. Café-s ‘cafés’) or unusual-sounding nonce words (e.g. Fnöhk-s), acronyms (e.g. BMW-s ‘BMWs’) and truncations (e.g. Sozi-s ‘socialists’) (see also Pinker & Ullman, 2002a: 459). Thus the class of regular, rule-bound forms does not have to be the largest in order to serve as the default form (minority default argument).

More recently, weaker versions of the dual-route model have been proposed which argue that the German plural system has several regular plural formations, such as the -s plural and the -(e)n plural for feminines ending in schwa (Wiese, 1999; Sonnenstuhl & Huth, 2002; see also Clahsen, 2004: 13f.).

A second view of German pluralization adopts the single-route model and proposes that -s is not qualitatively different from all other pluralization mechanisms (e.g. Bybee, 1995; Köpcke, 1998; Hahn & Nakisa, 2000). Instead, it is argued that -s is another case of the graded, probabilistic and context-sensitive operations typical of connectionist networks. Thus proponents of the single-route account stress that the German -s plural does not apply uniformly as the symbolic rule account predicts. According to Hahn & Nakisa (2000: 346f.), the only case of high and nearly uniform use of -s occurs with surnames and does not extend fully even to first names: two members of the Mann family are called Mann-s but two girls named Ulrike can be two Ulrike-n. Furthermore, they note a relatively high probability of using -s for foreign borrowings ending in full vowel, a specificity which undercuts the notion that -s is in any sense a default (Bybee, 1995: 442; see also McClelland & Patterson, 2002: 469).

Acquisition studies on German noun plurals
Psycholinguistic evidence for the acquisition of German noun plurals cited as support for the dual-route model comes primarily from experiments with nonce words and from studies on overgeneralization errors observed with German-speaking children.

In an acceptability judgment experiment, Marcus et al. (1995) asked adult German speakers to rate plurals of nonce nouns for acceptability. Results indicated that ‘irregular plural forms’ (i.e. all plural markers except for -s) were judged as significantly better when rhyming with actual German nouns (e.g. Pund, see Hund ‘dog’) than in a non-rhyme condition; -s plurals, by contrast, were judged as significantly better in the non-rhyme condition than in the rhyme condition. Based on this study, Bartke, Marcus & Clahsen (1995) obtained similar results with German-speaking children (age range 3;1–8;11).
The authors take these findings as evidence for the view that children (like adults) possess two distinct mechanisms for German plural inflection—a symbol-manipulating rule system for ‘regulars’ (-s plurals) and lexical entries for ‘irregulars’ (all other plural markers): only when the child fails to get access to a lexical entry for an irregular form (which is the case for unusual-sounding words such as non-rhymes and for non-canonical roots such as proper names), is the regular -s plural rule applied. Due to this qualitatively different status of -s as opposed to all other German plural markers, German speakers generalize -s widely, despite its low type frequency (Marcus, 2000: 163).

So far however, all studies on overgeneralization errors observed with German-speaking children have found that children overgeneralize a number of different plural markers; that the -s plural is never the only plural marker that occurs in overgeneralizations; and that -s is not even the most frequent plural marker occurring in overgeneralizations (see most recently Bittner & Köpcke, 2001; Schaner-Wolles, 2001; Szagun, 2001; Behrens, 2002) – counter to Clahsen et al.’s (1992: 236–8, 1996) claim.

A number of researchers regard the diversity of plural error patterns observed in German-speaking children as evidence for frequency and analogy-based generalizations, and thus do not assign a qualitatively different status to the German -s plural.

Köpcke (1998) adopts an output-oriented schema-based approach towards the acquisition of German plurals. Based on data from nonce-word tasks with German-speaking children (age range 6;0–7;0) and on longitudinal spontaneous speech data of seven German children (age range 1;11–2;9), he argues that cue strength (i.e. salience, type frequency, cue validity and iconicity) of plural markers plays a key role in the acquisition of the German plural morphology; and that overgeneralization errors are characterized by a drift from less reliable towards more reliable plural schemas. Bittner & Köpcke (2001) extend this analysis to propose a theoretical integration between the schema-model and naturalness criteria (e.g. transparency, uniformity and constructional iconicity) elaborated within natural morphology.

Based on dense longitudinal data of one German boy (aged 2;0–2;6), Behrens (2002) also provides counter-evidence against the assumption of the qualitatively different (i.e. default) status of the German -s plural. In line with the schema-based approach, she shows that -s is not the dominant marker erroneously used by the child, and that, contrary to the default assumption, -s overgeneralization errors are limited to distributional conditions present in the input to the child.

Finally, in her analysis of spontaneous speech corpora of 22 German-speaking children (age range 1;4–3;8), Szagun (2001) shows that the children’s growth rates of type frequencies per different plural marker
correspond to frequencies observed in the input addressed to them, and that
distributional information plays an important role in the children’s use of
error patterns.

Confounding factors in explaining German noun plurals
Most current views on noun plurals in German suffer from a number of
linguistic shortcomings (see Dressler, 1999; Wegener, 1999). Specifically,
they confound a number of notions.

Regularity vs. generality. First, the regular/irregular dichotomy confounds
regularity with generality of application. Regularity means, in a strict sense,
accountability by a symbolic rule or a recurrent pattern. In that sense, both
English verbs work – worked and sing – sang – sung are regular, however they
differ in generality of application and thus in type frequency (the more
general the application is, the more types get involved). Note that children
systematically overgeneralize the non-general pattern to bring – brang –
brung, but they very rarely overgeneralize an irregular pattern such as
bring – brought to sing – sought.

Regularity vs. default. Second, the notion of regularity is confounded
with the notion of default. Both Clahsen and Pinker (1999: 214, 222)
compare the status of the German -s plural to that of the English -s plural.
But the English -s plural is the only productive plural marker and
represents the overall default, whereas the German counterpart is only the
exceptional default or ‘emergency plural ending’ (Wegener, 1999): the -s
plural is used in special circumstances (see above) such as with names, in
word class changes or with newly loaned foreign words, e.g. Russian dača
‘country house’ > German Pl. die Datscha-s (later also Datsch-en). And
there are additional defaults as well: the -(e)n plural represents the default
for feminine nouns, and, for masculines and neuters, the default is the -e
suffix. Note, for example, the plurals die Quizz-e, die Fax-e of the masculine
and neuter English loan words der Quiz and das Fax, and the plurals die
Box-en, die Miss-en of the feminine English loan words die Box and die
Miss. Here, the final sibilant in masculine/neuter [kvis, faks] and feminine
[boks, mis] excludes an -s plural. Instead, the masculine/neuter default
plural suffix -e must be applied to masculines and neuters, the feminine
default plural suffix -e(n) to feminines. -s is also excluded after final
schwa in common nouns. Moreover, inflection of proper names cannot
provide evidence for default status, because inflection of names may
differ from inflection of common nouns in idiosyncratic ways in many
languages.

Schemas vs. inflection classes. Third, most psycholinguistic studies of
German declension assume that the plural formation is treated as a set of
rules (or schemas) distinct from the rest of the system of noun declension.
In reality, the choice of plural formation depends largely on gender and/or inflection class as manifested also in the expression of the four German cases in the singular. Thus if a masculine has the suffix -\textit{en} in the Gen.Sg., it must also have it for the plural, e.g. \textit{der Fürst} ‘prince, sovereign’, Gen.Sg. \textit{des Fürst-en} is a pattern which implies the plural \textit{Fürst-en}.

\textit{Morphological productivity}

Following Schultink (1961), we would like to propose that German noun plurals – and inflectional formation in general – operate under the constraints of \textit{morphological productivity}, defined as language-users’ ability to freely coin different morphological forms. Morphological productivity is inherent to all the approaches described above, from both single- and dual-route perspectives.

Under a dual-route view, only ‘regulars’ are fully productive (excluding degrees of productivity). Lower degrees of productivity can be assigned to ‘irregulars’ due to analogical expansion, particularly via rhyming words (e.g. \textit{bring – brang} after \textit{sing – sang}) or some other (undefined!) pattern similarity. One exception is the extended dual mechanism approach proposed by Sonnenstuhl & Huth (2002) which identifies four different degrees of predictability (or productivity) for the German -(\textit{e})\textit{n} plural and proposes to account for it by three types of inflection: (1) regular default, (2) class-specific inflection and (3) irregular inflection (pp. 279, 287).

Under a single-route view, differences in degree of productivity are accounted for by differences in frequency and (phonological or semantic) similarity. One single-route model which has focused on the role of productivity is Bybee’s network model (1995). According to Bybee (1995: 438), two main factors determine the productivity of a morphological pattern: (1) its type frequency and (2) the extent to which the defining schema of the pattern is open to a wide variety of phonological types.

\textit{A productivity scale for German noun plurals: a proposal}

We consider morphological productivity as \textit{the ability to form new potential words}. For the plural system, this means the ability to form plural forms of new words and potential variants of existing plural forms (Dressler, 2004).

We propose that productivity works on a broader scale than the seven categories with different plural markers presented in Table 1. \textit{Plural markers} are suffixes (or a zero morpheme) and stem vowel change (Umlaut). Productivity operates on \textit{plural patterns} which comprise three defining elements: (1) type of plural marker (Table 1 above), (2) gender (feminine vs. masculine/neuter, as signalled in the singular by articles, other
determiners and anaphoric pronouns), and (3) word-final phonology of the singular noun.

Productivity is closely related to predictability of morphological patterns: the more productive a plural pattern, the more predictable the plural output of a singular input which fits the structural description of the pattern.

Our productivity scale first makes a distinction between those plural patterns which apply freely to new words and are therefore productive, and those which practically do not occur with new words and are therefore non-productive. Productive patterns are (with the restrictions mentioned above in the discussion of regularity): -s plurals, -(e)n plurals of feminine nouns and of those masculine nouns which end in schwa, -e plurals (with and without umlaut) of masculines and neuters, and zero plurals of nouns ending in schwa plus sonorant. Non-productive patterns are notably a combination of the above-mentioned plural markers with other than the specified genders and word-final phonology, pure Umlaut plurals, and -er plurals (with or without Umlaut). Second, our productivity scale distinguishes three degrees within the productive patterns: (a) fully productive, (b) productive, and (c) weakly productive plural patterns, according to degree of competition with other productive plural patterns (see also Libben, Jarema, Dressler, Stark & Pons, 2002). Note that only competition between productive plural patterns is accounted for. Non-productive patterns are not qualified as being able to successfully compete with productive plural patterns. This yields the following four degrees of productivity within categories with different German plural markers: full productivity, productivity, weak productivity, and non-productivity.

The degree of fully productive is assigned to a productive plural pattern without any effective competing productive pattern. Full productivity implies maximal application to new words, such as loan words and abbreviations. For example, the integration of the French masculine noun *le garage* [garaʒ] into German feminine Sg. *die Garage* [ʒaʁəɡaː] – Pl. *die Garage-n* ‘garage’ shows that -(e)n plural is fully productive with German feminine nouns ending in schwa. In fact, all feminine and animate masculine nouns ending in schwa take the -(e)n plural, and there are thousands of such new words, including nouns derived from verbs, e.g. *liegen* ‘to lie’ > fem. Sg. *die Liege* – Pl. *die Liege-n* ‘couch’.

The degree of productive is assigned to a productive plural pattern with one other productive pattern as competitor for applying to the same base. For example, -(e)n plural formation to feminine nouns which end in consonants competes with -s plural formation: Sg. *die Farm* – Pl. *die Farm-en/Farm-s* ‘farm’.

The degree of weakly productive is assigned to a productive plural pattern with two other competing productive patterns. For example, the
plurals of the loan word *der Park* ‘park’ are *die Park-s*, *die Park-e*, and *die Pärk-e* (the last one actually used in Switzerland and a potential variant elsewhere, see Wegener, 1999: 2).

The degree of non-productive is assigned to plural patterns which practically do not occur with new words. For example, the pure Umlaut plural (e.g. *der Mantel* – *die Mäntel* ‘coat’), the -er plural, with or without Umlaut (e.g. *der Wurm* – *die Würm-er* ‘worm’, *das Bild* – *die Bild-er* ‘picture’) and the -(e)n plural with neuter nouns (e.g. *das Bett* – *die Bett-en* ‘bed’).

Table 2 presents an overview of the plural patterns and plural markers in German, with respect to their degree of productivity.

The effects of productivity interact with type and token frequency. Type frequency of a morphological pattern is, to a certain extent, conditioned by productivity: the more productive an inflectional pattern, the more different forms are expected to occur, since a more productive pattern applies to more words than a less productive pattern. Overall token frequency of inflectional patterns is in turn positively influenced by greater type frequency and thus indirectly also by productivity. One example for the difference between productivity and frequency is the German -s plural: the -s plural is productive for all three genders, is potentially applicable to the greatest segment of new German nouns either as the only possibility or in competition with other plural markers and is thus the most productive German plural formation. But actually its type and token frequency is relatively low, because it applies to relatively few old words (Goebel & Indefrey, 2000: 180).

The relation between lexical frequency and morphological productivity in acquisition is of crucial importance theoretically, because single-route and dual-route models make divergent predictions on this issue. Under a single-route approach, one would expect to find lexical frequency effects for all German plural markers, including -s plurals. Under a dual-route approach, one would expect to find lexical frequency effects for irregular plurals, whereas no such effect should be observed for -s plurals, especially for those in a fully productive environment. Note that if one adopts the parallel-race model variant of the dual route model proposed by Pinker & Ullman (2002a: 458; see also section models of processing), only the -s plurals classified as fully productive in our productivity scale would be generated by rule, because the fully productive environments for -s are environments with unattested or at best very small families of irregulars; all -s plurals classified as productive or weakly productive would be stored, because they resemble (and thus compete with) large or very large families of irregulars.

Productivity also interacts with morphological transparency (and hence with salience, Köpcke, 1998). German pluralization is transparent
<table>
<thead>
<tr>
<th>Plural pattern</th>
<th>Productivity</th>
<th>Examples (test items)</th>
<th>Examples (loan words)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>singular</td>
<td>plural</td>
</tr>
<tr>
<td>-s* -a</td>
<td>fully productive</td>
<td>—</td>
<td>der Pyjama</td>
</tr>
<tr>
<td></td>
<td>productive</td>
<td>-(e)n</td>
<td>die Pizza</td>
</tr>
<tr>
<td>other full vowel</td>
<td>productive</td>
<td>—</td>
<td>das Baby</td>
</tr>
<tr>
<td>schwa plus n/r</td>
<td>weakly productive</td>
<td>-(e)n, zero</td>
<td>der/das Poster</td>
</tr>
<tr>
<td>shwa plus l</td>
<td>weakly productive</td>
<td>-e, -e + U</td>
<td>der Clown</td>
</tr>
<tr>
<td>n/r/l after non-schwa; other consonants</td>
<td>weakly productive</td>
<td></td>
<td>der Hase</td>
</tr>
<tr>
<td>other conditions</td>
<td>non-productive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-(e)n schwa</td>
<td>fully productive</td>
<td>—</td>
<td>der Katze</td>
</tr>
<tr>
<td>(Continued)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Continued)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### EARLY NOUN PLURALS IN GERMAN

<table>
<thead>
<tr>
<th>Gender</th>
<th>Endings</th>
<th>Description</th>
<th>Example 1</th>
<th>Example 2</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>fem.</td>
<td>-s</td>
<td>productive</td>
<td>die Uhr</td>
<td>die Uhr-en</td>
<td>see <em>Diva</em> above</td>
</tr>
<tr>
<td></td>
<td>-s, zero</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>masc., neut.</td>
<td>-s</td>
<td>productive</td>
<td>das Bett</td>
<td>die Bett-en</td>
<td>see <em>Pixel</em> above</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>consonant</td>
<td>-s§</td>
<td>non-productive</td>
<td>der Zug</td>
<td>die Züg-e</td>
<td>see <em>Park</em> above</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other conditions</td>
<td>-e, -s¶</td>
<td>non-productive</td>
<td>der Pullover</td>
<td>die Pullover</td>
<td>see <em>Poster</em> above</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>schwa plus l</td>
<td>-e + U</td>
<td>non-productive</td>
<td>der Mantel</td>
<td>die Müntel</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>zero</td>
<td>-(e)n, -s</td>
<td>productive</td>
<td>der Wurm</td>
<td>die Würm-er</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>-er + U</td>
<td>non-productive</td>
<td>der Mantel</td>
<td>die Müntel</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Never after sibilants.
† Nouns of Latinate or Slavic origin; other, rarer Latinate patterns are excluded from the table.
‡ No competition with -e + U in the case of a non-umlautable stem vowel, i.e. productive instead of weakly productive.
§ No competition with -s in the case of a word-final sibilant, i.e. fully productive instead of productive.
¶ No competition with -e in the case of a word-final sibilant, and no competition with -e + U in the case of a non-umlautable stem vowel, i.e. productive instead of weakly productive; when competition neither with -s, nor with -e + U: fully productive.
• No competition with -s in the case of a word-final sibilant, i.e. productive instead of weakly productive.
where there is either no change in the plural (zero plural), e.g. Sg. *Pullover* ‘pullover’, Pl. *Pullover* or a suffix is added without any further change, e.g. Sg. *Baby* ‘baby’, Pl. *Baby-s*; it is opaque when the singular base is changed via umlaut, as in Sg. *Mantel* ‘coat’, Pl. *Mäntel* or Sg. *Wurm* ‘worm’, Pl. *Würm-er*. Thus, all fully productive patterns are transparent, whereas the entirely non-productive pure Umlaut is opaque (*Mäntel*). However, non-productive -er plurals may be either transparent (e.g. *Bild-er* ‘pictures’) or opaque (*Würm-er*). The same holds for weakly productive -e plurals: transparent Pl. *Schaf-e* ‘sheep’ vs. opaque Pl. *Bäll-e* ‘balls’ or *Generäl-e* ‘generals’ (earlier on transparent *General-e* and still earlier *General-s*).

**Predictions**

We hypothesized that degree of productivity is one relevant predictor for the probability of children’s correct production of plural patterns. In case of plural pattern competition, the child must choose between two or three possible patterns. Thus, correctness of production in the current study should pattern according to our productivity scale.

Degree of productivity should also play a role in the distribution of error types, so that more productive plural patterns should attract more errors with less productive target patterns (error direction from less to more productive).

For correct responses, we predicted the highest scores for fully productive plural patterns (e.g. *Baby-s*), less high scores for productive patterns (e.g. *Pizza-s*) and still lower scores for weakly productive patterns (e.g. *Clown-s*). For erroneous responses, we predicted a similar pattern of distribution, with fewest errors in the direction of non-productive patterns.

**THE AUSTRIAN GERMAN NOUN PLURAL STUDY**

To test our proposal, we designed an experimental study, which examines how Viennese children learn about noun plural formation.

**METHOD**

**Participants**

84 children (44 boys and 40 girls) aged 2;6 to 6;0 took part in this study. Participants were grouped into seven different age groups in intervals of 6 months, consisting of 12 children each: two-year-olds (age range 2;6–3;0, mean age 2;10); younger three-year-olds (age range 3;1–3;6, mean age 3;4); older three-year-olds (age range 3;7–4;0, mean age 3;10); younger four-year-olds (age range 4;1–4;6, mean age 4;4); older four-year-olds (age range 4;7–5;0, mean age 4;10); younger five-year-olds (age range 5;1–5;6,
mean age 5;4); older five-year-olds (age range 5;7–6;0, mean age 5;10). All children were monolingual speakers of Austrian German, with no developmental or linguistic problems, and attended private kindergartens in a middle/high socioeconomic neighbourhood in Vienna.

Materials and procedure
The study was an experimental design in which participants were asked to give the plural form of 42 stimulus items. Items were balanced for the seven plural markers presented in Table 1 and for item gender (if applicable). For the complete list of test items, see the Appendix.

Items were selected with regard to input token frequencies measured in the longitudinal spontaneous speech corpus of one Austrian girl (Katharina, see Vollmann, Sedlak, Müller & Vassilakou, 1997). In order to exclude clearly rote-learned plural forms from the test, we focused on nouns with middle input token frequency (N=5–10). Clearly plural-oriented items (e.g. Ei-er ‘eggs’), singular-oriented items (e.g. Mund ‘mouth’), and abstract or non-imageable items (e.g. Lied ‘song’, Rand ‘margin’) were excluded. In cases in which the above input criteria could not be fulfilled, slightly more/less frequent items were added (e.g. Loch ‘hole’: N=11, Kuh ‘cow’: N=4). Finally, missing items were integrated by using frequency information from the German CELEX database (this was especially the case for the pure Umlaut marker, where we had to choose the items Apfel ‘apple’ (CELEX: N=74) and Vogel ‘bird’ (CELEX: N=235), although they were highly frequent in our input data (N=32 and 21, respectively) and to add the items Hammer ‘hammer’ (CELEX: N=74) and Nagel ‘nail’ (CELEX: N=74), although they did not occur at all in our input data).

Children were tested orally and individually in their kindergarten. Plural forms were elicited in the following way: the child was presented with a picture depicting a singular noun (e.g. Hase ‘rabbit’), and the investigator said: Das ist ein Hase ‘This is a rabbit’. Then, a second picture depicting three instances of the same noun was shown to the child, and the investigator asked: Und was sind das? Das sind drei/viele ______ ‘And what are these? These are three/many ______’. Test items were presented in two random orders and were preceded by three training items (Auto – Auto-s ‘car-s’, Banane – Banane-n ‘banana-s’, Baum – Bäum-e ‘tree-s’).

Data coding and the problem of zero plurals
Test items were coded for plural marker (-s, -(e)n, -e, -e + U, zero, U, -er +U), item gender (masc., fem., neut.), degree of productivity (thus distinguishing between fully productive, productive and weakly productive -s plurals; fully productive, productive and non-productive -(e)n plurals, etc.),
and token frequency (high, middle, low) in two different measures: input and CELEX (see Appendix).

An especially problematic point (often neglected in prior investigations of plural formation in German) was the demarcation of singular forms vs. incorrect zero plurals in German test data. How can be accounted for the fact that a zero form such as Mantel ‘coat’ (for correct plural Mántel ‘coats’) might be considered either a mere repetition of the given singular stimulus, or else an overgeneralization of the zero plural marker? For the purpose of this study, we decided to code zero forms as incorrect zero plurals when at least two of the following three criteria were satisfied: (1) more than 50% of the child’s responses had overt (though not necessarily correct) plural marking; (2) zero forms were interspersed with overtly marked plural forms (i.e. not more than 5 zero forms in succession); and (3) the child occasionally ‘managed’ the plural task, e.g. by taking over the role of the experimenter or by using plural quantifiers (Brown, 1973: 331). In cases where only one or none of the three criteria was fulfilled, zero forms were coded as repetition of singular forms. Another possible solution to the problem of zero plurals would have been to use a testing procedure which avoids priming the zero-marked form (see, for example, the procedure used by Stemberger & MacWhinney, 1986: 22 for eliciting English past tense).

The distribution of satisfied criteria per age group is given in Table 3. As can be seen, the simultaneous satisfaction of at least two criteria is at 58% in the youngest age group. Beginning with the group of older three-year-olds, it is above 90%.

**RESULTS**

Participants’ responses were classified into either correct (i.e. corresponding to plural forms used by adults speaking Standard Austrian German) or incorrect. We first present the results for the correct responses.
Correct responses

Table 4 shows the means of correct responses for the six different overt plural markers (-s, -(e)n, -e, -e + U, U, -er + U), by age group. The maximum number of correct responses per plural marker was 6.

As can be seen, correct responses increase with age. Highest correct scores were obtained with -e plurals \( (m=3.80) \), followed by the plural markers -e + U \( (m=3.06) \), -er + U \( (m=2.54) \), -s \( (m=2.40) \), and -(e)n \( (m=2.26) \). The lowest score was observed for pure Umlaut plurals \( (m=1.02) \). Note that, in contrast to incorrect zero forms (see section data coding), all correct zero forms have been coded as correct zero plurals. Thus, it can not be excluded that some of these plural forms are mere repetitions of the singular form elicited by the investigator. For this reason, zero plurals have been excluded from the analysis of correct responses.

Based on the data presented in Table 4, we carried out a five-factorial two-way ANOVA combining the factors (7) age \( \times (6) \) plural marker \( \times (3) \) item gender \( \times (4) \) productivity \( \times (3) \) frequency. Frequency was tested twice: in a first run for input frequency, and in a second run for CELEX frequency. In order to approximate normal distribution and equality of variance of the data, individuals have been grouped together by age groups. This analysis showed significant main effects for age \( (F(6,147)=10.70, p<0.001) \), plural marker \( (F(4,147)=10.59, p<0.001) \) and productivity \( (F(3,147)=6.58, p<0.001) \), but not for item gender nor for either type of the frequency values (input and CELEX). The observed main effects for age indicate that there is a rise with age in the correct responses. Pairwise comparisons of the age groups at the 0.001 level (Bonferroni) showed that the two youngest age groups \( (2;6–3;0 \text{ and } 3;1–3;6) \) differ from the two oldest ones \( (5;1–5;6 \text{ and } 5;7–6;0) \). The observed main effects for plural marker indicate that the six plural markers differ in their correct scores. Pairwise comparisons of -e plurals and pure Umlaut plurals at the 0.001 level (Bonferroni) showed that -e plurals differ from all plural markers except for -e + U plurals, and that pure Umlaut plurals differ from all other plural markers. Furthermore, the results obtained indicate that productivity has an impact on the number of correct scores; and that productivity is an independent factor from frequency. The lack of an overall frequency effect might be due to the fact that frequency showed stronger effects for some plural markers than for others (see the more specific analyses below).

The distribution of correct responses with respect to productivity is given in Figure 1: fully productive and productive plural patterns showed higher correct scores than weakly productive and non-productive ones. Pairwise comparisons of the four degrees of productivity at the 0.001 level (Bonferroni) showed that – except for fully productive and productive plural patterns – all degrees of productivity differ from each other.
**Table 4.** Means (m) and standard deviations (S.D.) of correct responses for the 6 different overt plural markers, by age group (maximum 6)

<table>
<thead>
<tr>
<th></th>
<th>2;6–3;0</th>
<th>3;1–3;6</th>
<th>3;7–4;0</th>
<th>4;1–4;6</th>
<th>4;7–5;0</th>
<th>5;1–5;6</th>
<th>5;7–6;0</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m</td>
<td>S.D.</td>
<td>m</td>
<td>S.D.</td>
<td>m</td>
<td>S.D.</td>
<td>m</td>
<td>S.D.</td>
</tr>
<tr>
<td>-(s)</td>
<td>1.00</td>
<td>0.95</td>
<td>2.17</td>
<td>1.34</td>
<td>3.25</td>
<td>1.76</td>
<td>2.08</td>
<td>1.93</td>
</tr>
<tr>
<td>-(e)n</td>
<td>1.83</td>
<td>1.75</td>
<td>1.92</td>
<td>1.51</td>
<td>2.25</td>
<td>1.36</td>
<td>2.25</td>
<td>1.48</td>
</tr>
<tr>
<td>-(e)</td>
<td>2.42</td>
<td>1.98</td>
<td>3.67</td>
<td>1.72</td>
<td>4.58</td>
<td>1.38</td>
<td>3.42</td>
<td>0.79</td>
</tr>
<tr>
<td>-(e + U)</td>
<td>2.00</td>
<td>1.54</td>
<td>2.08</td>
<td>1.83</td>
<td>2.83</td>
<td>2.21</td>
<td>3.00</td>
<td>1.60</td>
</tr>
<tr>
<td>U</td>
<td>0.83</td>
<td>1.19</td>
<td>0.58</td>
<td>1.16</td>
<td>1.00</td>
<td>1.35</td>
<td>0.75</td>
<td>1.48</td>
</tr>
<tr>
<td>-(e + U)</td>
<td>1.42</td>
<td>1.51</td>
<td>1.58</td>
<td>1.83</td>
<td>1.75</td>
<td>1.86</td>
<td>2.83</td>
<td>1.53</td>
</tr>
<tr>
<td>Total</td>
<td>2.02</td>
<td>1.90</td>
<td>2.25</td>
<td>1.85</td>
<td>2.74</td>
<td>2.01</td>
<td>2.76</td>
<td>1.87</td>
</tr>
</tbody>
</table>
There was a significant interaction between plural marker and item gender ($F(1, 147) = 6.54, p = 0.012$) which indicates that the three genders show different correct scores depending on the plural marker. An overview of this interaction is given in Figure 2. Interrupted lines indicate that the respective gender was not applicable for the observed plural marker.
However, there was no significant interaction between age and plural marker, nor between age and productivity which indicates that type of plural marker and degree of productivity played an equal role in all age groups.

In order to determine whether the six overt plural markers differ with respect to the impact of age, item gender, productivity and frequency, we carried out a second analysis in which each plural marker was tested separately. Due to the reduced number of data available within each plural marker, we conducted a four-factorial two-way Logit-ANOVA: (7) age × (3) item gender × (4) productivity × (3) frequency; individuals have been grouped together by age group. The following sections present a detailed analysis of correct responses in each plural marker by descending order of correct scores. When all items except for one had the same value (see Appendix), we did not test for productivity (-er + U and pure Umlaut plurals), for input frequency (-e plurals), for CELEX frequency (-e + U and -s plurals), or for item gender (pure Umlaut plurals).

The following figures show, for each of the six plural markers, the means of correct responses by productivity (Figure 3), by input frequency (Figure 4) and by CELEX frequency (Figure 5).

-e plurals. The -e plurals are the most frequent overt plural marker in children’s correct responses. Due to extreme heteroscedasticity (i.e. inhomogeneity of variance), no adequate logit model could be fitted to the data. Results obtained by descriptive statistics point to possible influences of age and productivity on the number of correct scores. Correct -e plurals rise steeply in the youngest age groups (Table 4). Correct scores were higher with the fully productive -e pattern Fisch-e ‘fish’ than with productive (e.g. Schiff-e ‘ships’) and weakly productive plural patterns (e.g. Schaf-e ‘sheep’) (Figure 3). However, there were no gender differences (Figure 2).

-e + U plurals. Within -e + U plurals, only age showed a significant effect (p=0.006), indicating that there is an increase of correct -e + U plurals with age (Table 4). Item gender, productivity and input frequency were not significant.

[1] A logit model (Weisberg, 1985: 267f.) is the appropriate model to investigate the influence of factors on a categorial dependent variable which is given by the odds-ratio (nb of correct answers/nb of total answers). Using the logit-transformed odds-ratios leads to a continuous dependent variable which can be modelled as a linear combination of several categorial factors. This model can be investigated by analysis of deviance (see Hastie & Pregibon, 1993: 195f.) – a statistical test of the influence of factors and their interactions on a dependent variable. Because of this analogy to the ANOVA, we will call this analysis a logit-ANOVA. Using an ordinary ANOVA on a categorial dependent variable might be dangerous, insofar as it may lead to unjustified significances due to violated assumptions.
-er + U plurals. Analysis of correct responses for the -er + U plurals yielded significant main effects for age ($p < 0.001$) and item gender ($p < 0.001$): there was a rise of correct -er + U plurals with age (Table 4) and neuter nouns had higher scores than masculines (Figure 2). CELEX frequency (Figure 5) was significant ($p = 0.005$): high frequency items
scored higher than middle and low frequency items. Input frequency was not significant.

*s plurals. Significant effects were observed for all investigated factors: age \((p < 0.001)\), item gender \((p < 0.001)\), productivity \((p < 0.001)\) and input frequency \((p < 0.001)\). Table 4 shows that correct scores for *s plurals rise with age, especially in the youngest age groups. Neuters and feminines have higher scores than masculines (Figure 2). Figure 3 shows that fully productive -s patterns (e.g. Baby-s ‘babies’) scored higher than productive (Pizza-s ‘pizzas’) and weakly productive (Clown-s ‘clowns’) -s patterns. High and middle input frequency items were easier than low frequency items (Figure 4).

A Kruskal-Wallis test restricted to the group of fully productive -s plurals showed also effects for input frequency \((p < 0.001)\): middle and high frequency items scored higher than low frequency items (see Table 5).
-\(e\)n plurals. Analysis of correct scores showed significant main effects for item gender \((p < 0.001)\) and productivity \((p < 0.001)\), but not for age (Table 4). CELEX frequency was significant \((p = 0.010)\), while input frequency was not. Figure 3 shows that fully productive -\(e\)n patterns (e.g. Katze-n ‘cats’) and productive patterns \((Uhr-en ‘watches’)\) showed high correct scores from early on, but non-productive -\(e\)n patterns \((Hemd-en ‘shirts’)\) remained difficult even for the oldest age groups. High frequency items showed more correct scores than middle and low frequency items (Figure 5). Concomitantly, feminine items had the highest scores, followed by masculines and neuters (Figure 2).

Pure Umlaut plurals. This plural marker had the lowest number of correct scores. Analysis yielded a significant main effect for age \((p = 0.016)\): Table 4 above shows that U plurals slightly increase with age. Input frequency (Figure 4) was significant \((p < 0.001)\), but CELEX frequency was not. Most correct scores were observed for high frequency items (e.g. Vogel ‘bird’), followed by middle (Mantel ‘coat’) and low frequency items (Hammer ‘hammer’).

Incorrect responses
Incorrect responses were classified into three major groups: (i) repetitions of singular forms; (ii) overgeneralizations; and (iii) other errors (i.e. no response, use of another item, use of another form, e.g. diminutive marking). Overgeneralizations were further divided into nine overgeneralization types:

1. \(-s\), e.g. Tiger – *Tiger-s (← Tiger ‘tigers’)
2. \(\text{\textit{-}(e)n}\), e.g. Pizza – *Pizza-n (← Pizza-s ‘pizzas’)
3. \(-e\), e.g. Hemd – *Hemd-e (← Hemd-en ‘shirts’)
4. \(-e + U\), e.g. Bus – *Büss-e (← Buss-e ‘buses’)
5. \(\text{\textit{zero}}\), e.g. Mantel – *Mantel (← Mäntel ‘coats’)
6. \(U\), e.g. Pyjama – *Pyja¨ma (← Pyjama-s ‘pyjamas’)
7. \(-\text{\textit{er}}\), e.g. Bett – *Bett-er (← Bett-en ‘beds’)
8. \(-\text{\textit{er}} + U\), e.g. Zug – *Züg-er (← Züg-e ‘trains’)
9. double marking, e.g. Katze – *Katze-n-s (← Katze-n ‘cats’)

i. Repetitions of singular forms make up 15% of all incorrect responses; this group of errors was predominant in the two youngest age groups (39% and 34% of all incorrect responses respectively).

ii. Overgeneralizations constitute 78% of the incorrect responses. Within overgeneralizations, \(\text{\textit{zero}}\) forms were most frequent \((N = 583, 31\%)\). Note however that this result has to be interpreted very cautiously, since the demarcation of singular forms vs. incorrect \(\text{\textit{zero}}\) plurals drawn for our analysis (section data coding) is still – to a certain extent – an arbitrary
distinction. While it is rather unlikely that zero forms observed with children not fulfilling the three criteria are incorrect zero plurals, it can not be excluded that zero forms observed with children fulfilling them are (at least sometimes) singulars. Within overtly marked overgeneralizations, -e suffixation was the most frequent error type ($N=312$, 16%), followed by -(e)n ($N=197$, 10%), -s ($N=184$, 10%) and -e + U ($N=126$, 7%). All other plural markers occurred very rarely in overgeneralizations ($U$: $N=7$, 0%; -er: $N=29$, 2%; -er + U: $N=19$, 1%; double marking: $N=24$, 1%). Analysis of overtly marked overgeneralizations by age group revealed two peaks of overgeneralizations, a first one in the group of older three-year-olds ($N=171$ out of 504 possible responses, 34%) and a second, (smaller) one in the group of younger five-year-olds ($N=149$, 30%); in the oldest age group, overt overgeneralizations constitute still 24% ($N=120$) of all possible responses. In order to see whether age has an impact on the type of overgeneralizations used, we conducted a one-way Logit-ANOVA ((7) age $\times$ (1) overgeneralization type) on the five most frequent overgeneralization types: zero, -e, -(e)n, -s and -e + U; individuals have been grouped together by age group. This analysis showed a significant effect for the -e + U marker ($p=0.004$), but not for zero, -e, -(e)n, and -s. These results indicate that overgeneralizations of older children differ from those of younger children by an increased use of Umlaut in -e plurals: i.e. overgeneralization errors such as Cläun-e (← Clown-s ‘clowns’), Büb-e (← Bub-en ‘boys’), Büss-e (← Buss-e ‘buses’) appear to be typical of older children.

Analysis of error direction showed that the majority of all overtly marked overgeneralizations (78%, $N=703$ out of 898) result either from shifts from less productive plural patterns towards more productive ones or from competition between plural patterns of the same degree of productivity: overgeneralization of the suffix -e occurred most frequently in neuter -(e)n plurals, i.e. in cases where -(e)n suffixation is non-productive while -e plural formation is productive (e.g. *Hemd-e ← Hemd-en ‘shirts’). Incorrect use of -e was also rather frequent in the non-productive -er + U plurals (e.g. *Wald-e ← Wald-er ‘woods’), and in the competing -e + U plurals (e.g. *Zug-e ← Züg-e ‘trains’). Overgeneralization of -(e)n was observed most often with nouns of the non-productive pure Umlaut plural (e.g. *Nagel-n ← Nägel ‘nails’). Overgeneralization of -s was most frequent with zero plurals (e.g. *Pullover-s ← Pullover ‘pullovers’) and with the non-productive pure Umlaut plurals (e.g. *Hammer-s ← Hämmer ‘hammers’). Overgeneralization of -e + U occurred most often with the competing -e plurals (e.g. *Büss-e ← Buss-e ‘buses’) and with the non-productive -er + U plurals (e.g. *Würm-e ← Würm-er ‘worms’). Overgeneralization of the non-productive -er (e.g. *Bett-er ← Bett-en ‘beds’), -er + U (e.g. *Hüt-er ← Hüt-e ‘hats’), and pure U plurals (e.g. *Pyjäma ← Pyjama-s ‘pyjamas’) occurred very rarely.
In addition to productivity, morphological transparency plays a relevant role in error direction: 41% ($N=371$) of the children’s overt overgeneralizations are more transparent than the target plural form (e.g. omission of opacifying Umlaut in *Zug-e ← Züg-e ‘trains’), 48% ($N=428$) are equally transparent (e.g. erroneous suffix use in *Hemd-e ← Hemd-en ‘shirts’), and only 11% ($N=99$) are less transparent. Note that the plural elicitation technique used in this study (see section materials and procedure) might have increased the production of transparent plurals, because the non-umlauted vowel (i.e. the singular form) has been primed.

In order to evaluate more precisely the role of productivity and morphological transparency for children’s overt overgeneralizations, we calculated the chance of independence of the two measures, by transparency of the target plural (see Table 6).

The analysis shows that the effects of transparency and productivity coincide for half of the errors ($T=P$: 50%), whereas they diverge for the other half: for 28% of the errors, productivity wins over transparency ($T<P$); for 22%, transparency wins over productivity ($T>P$). Analysis of transparent vs. opaque target plurals indicates that the chance of independence of the two measures differs by degree of transparency: for transparent target plurals, the chance of independence is at 62%, whereas for opaque plurals, it is only at 38%.

iii. 7% of the incorrect responses belong to the group of ‘other’ errors: in the youngest age group, most of these are either no response or use of another item (e.g. Vogel ‘bird’ instead of Schnabel ‘beak’); then, use of another form (e.g. Häschen ‘rabbit-DIM’ instead of Hase-n ‘rabbit-PL’) predominates.

**SUMMARY**

General analyses of correct responses showed that correct responses patterned with age, plural marker and productivity. Correct responses increased with age, especially between ages three and five. Within plural

**Table 6. Overt overgeneralizations: chance of independence ($T>P$; $T<P$) and non-independence ($T=P$) of transparency ($T$) and productivity ($P$), by transparency of the target plural**

<table>
<thead>
<tr>
<th>Transparency (target plural)</th>
<th>$T&gt;P$</th>
<th>$T=P$</th>
<th>$T&lt;P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transparent</td>
<td>18% (84/467)</td>
<td>38% (176/467)</td>
<td>44% (207/467)</td>
</tr>
<tr>
<td>Opaque</td>
<td>27% (116/431)</td>
<td>62% (269/431)</td>
<td>11% (46/431)</td>
</tr>
<tr>
<td>Total</td>
<td>22% (200/898)</td>
<td>50% (445/898)</td>
<td>28% (253/898)</td>
</tr>
</tbody>
</table>
markers, highest correct scores were obtained for -e plurals, followed by -e + U, -er + U, -s and -(e)n, with the lowest score observed for pure Umlaut (U) plurals (zero plurals were excluded from the analysis of correct responses). Correct production followed the predicted productivity scale fully productive > productive > weakly productive > non-productive. Neither type of the frequency values (input and CELEX) showed an overall significant effect.

Specific analyses of correct responses showed that correct scores increased for all plural markers, except for -(e)n. In analysing those plural markers containing items with different degrees of productivity, our productivity scale was supported separately for -s, -(e)n and -e, but not for -e + U plurals. Input frequency effects were observed for -s and U plurals, and CELEX frequency for -(e)n and -er + U plurals: in both cases high frequency items scored higher than middle and low frequency items.

Analyses of incorrect responses showed that -e suffixation was the most frequent overt error type, followed by -(e)n, -s and -e + U; all other plural markers occurred very rarely. In addition to productivity, morphological transparency was shown to play a relevant role in error direction.

**DISCUSSION**

This study examined the acquisition of noun plurals by 84 Viennese preschoolers aged 2;6 to 6;0, in an elicitation task in which children had to give the plural form of a given singular noun. Analyses focused on children’s production of correct responses. The following variables were investigated: age (7 age groups consisting of 12 children each), plural marker (-s, -(e)n, -e, -e + U, U, -er + U), item gender (masc., fem., neut.), degree of productivity (fully productive, productive, weakly productive, non-productive), and token frequency (high, middle, low) in two different measures: input and CELEX.

The most important finding of our study was the relevance of degrees of morphological productivity for children’s scores of correct responses: in accordance with our initial hypothesis, fully productive and productive plural patterns showed higher scores than weakly productive and non-productive ones.

In our analysis, we first made a distinction between plural patterns which apply freely to new words and are therefore productive and those plural patterns which do not occur with new words and are therefore non-productive. These different patterns were defined according to (1) plural marker, (2) gender and (3) word-final phonology. As a second step we distinguished three degrees within the productive patterns: (a) fully productive, (b) productive, and (c) weakly productive plural patterns, according to degree of competition with other productive plural patterns.
Thus we distinguished within our test items between fully productive, productive and non-productive -en plurals, fully productive, productive and weakly productive -s and -e plurals, etc. (see Appendix).

In contrast to productivity (i.e. the combined feature analysis of plural marker, gender and word-final phonology), a separate analysis of item gender did not show any overall significant effect on children’s correct scores: this result was expected, since all three genders are productive in German (as opposed to, e.g. Slovenian where the neuter gender in general is non-productive).

The basic role of degree of productivity was also apparent in the distribution of our children’s error patterns: the majority of all overtly marked overgeneralizations resulted either from shifts from less productive plural patterns towards more productive ones or from competition between plural patterns of the same degree of productivity (see also Klampfer(-Laaha), Korecky-Kröll & Dressler, 2001). The patterning of these errors testifies to children’s early grasp of the underpinnings of productive German plural morphology. In addition to productivity, morphological transparency (a factor of higher salience in Köpcke, 1998: 308) was shown to play a relevant role in error direction.

Degree of productivity played an equal role in all age groups, i.e. we found no interaction between age and productivity. This means that, already in the youngest age group of our study (2;6–3;0), children must have identified the relevance of the gender distinction between feminine and non-feminine and of word-final phonology. This indicates early sensitivity to typological features of the language, those features that are prominent and carry the most information – in line with similar sensitivity to typological features of Semitic languages such as roots and bound suffixes. These results fit findings on largely correct gender distinctions marked early on in the acquisition of the German article system (see Szagun, 2004). Children’s identification of pertinent conditions is, however, only partial, since correct responses never reach ceiling. Even if we take a level of 75% correct scores as satisfactory, this is only two times reached for -e plurals in older age brackets (see Table 4), but not at all for the other plural markers. Similarly, our analysis of incorrect responses showed no clear decrease of overt overgeneralizations in the oldest age group (5;7–6;0) – suggesting that at this age, learning of German noun plurals is still under way (see a similar situation in the acquisition of Arabic plurals, see Ravid & Farah, 1999).

How do our results fit dual- and single-route models? In general, our results do not provide evidence in support of dual-route models, but might be compatible with new, weaker versions of them – Pinker & Ullman’s parallel-race model (2002a) and Sonnenstuhl & Huth’s extended dual mechanism approach (2002).
Several of our findings are problematic for dual-route models. First, as explained in section morphological productivity above, under a classical dual-route approach, lexical frequency effects should be found only for irregular plurals, but not for -s plurals. Under the new race models, only fully productive -s plurals should show no such effects. In fact, we have found both kinds of effects of input token frequency, see section correct responses above. Secondly, in section models of processing above, we summarized Marcus et al.’s (1995) findings that irregular plural nouns were judged as better when rhyming with actual plural nouns than in a non-rhyme condition. Thus, the dual-route model predicts that rhymes would play a major role among overgeneralizations (see the concept of hyper-similarity in Marcus et al., 1995: 193–4). However, among the 714 examples of overt overgeneralizations (without -s), only 84 (12%) rhyme with plurals children may have heard (11 Clàun-e ‘clowns’ after Zàun-e ‘fences’, 8 Klô-e ‘toilets’ after Flôh-e ‘fleas’, 36 Büss-e ‘buses’ after Küss-e ‘kisses’, 16 Wûrm-e ‘worms’ after Tûrm-e ‘towers’, 13 Häus-e ‘houses’ after Mûus-e ‘mice’, but, e.g. no Mûus-er ‘mice’ after Häus-er ‘houses’, instead 10 examples of Maus-e). Thirdly, the dual-route view of German -s plurals representing a general ‘phonology-free’ default (Clahsen et al., 1992; Marcus et al., 1995; see section models of processing above) predicts that children overgeneralize -s plurals across the board. But, in fact, this systematic study of the acquisition of German plurals did not provide any such indications of a general ‘phonology-free default either in terms of the only or of the most frequently used plural marker in overgeneralizations: in our data, -s errors made up 20% of all overtly marked overgeneralizations and were less frequent than overgeneralizations with -e and -(e)n. Furthermore, overgeneralization of -s was restricted to specific phonological environments: it was – with the exception of very few examples (6 out of 184) – limited to nouns ending in a full vowel (or in a-schwa) (e.g. Kuh-s, Pullove[\']s), in a nasal (e.g. Mädche-n-s), or in a liquid (e.g. Nagel-s) (Klampfer(-Laaha) et al., 2001: 38; for similar findings see Szagun, 2001; Behrens, 2002).

Thus, two other approaches to the acquisition and processing of inflectional morphology are better suited to explain our findings. Our results are fully compatible with single-route models, which assume gradual differences – ranging from highly predictable to completely idiosyncratic plural patterns. In a single-route model, fully productive patterns (i.e. without competition) can be predicted to be handled best; the more competitors there are, the less success should be guaranteed. Least success can be predicted for non-productive patterns which have to compete, in our test items, with at least two productive patterns of various degrees of productivity (with the only exception of Maus – Mâuse ‘mice’ which competes with just one) and are no match to productive patterns. Morphological
transparency contributes to the success of more productive patterns (see section incorrect responses). Dual-route race-models as developed by Baayen & Schreuder (1999) might also accommodate our findings: after children have extracted rules, the more productive and transparent a rule, the more likely its application to plural patterns is to win over direct lexical access. Within this race model, we can interpret children’s failure to approach ceiling in correct responses as lack of routine in applying the rules they have extracted.

This paper argued for the importance of degrees of productivity in explaining patterns of morphological acquisition. We provided a two-step productivity scale to characterize such degrees of productivity, in our case for German plurals. In summary, the general tendencies observed in both correct and incorrect responses are very much in accordance with our productivity scale. Complex inflection systems such as German plural formation, number-and-case systems or verb-inflection systems of Slavic languages are not reducible to simple models of ‘regular’ vs. ‘irregular’ inflection. This holds both for linguistic description (including diachronic change) and for all types of psycholinguistic investigations, from acquisition over processing to language impairments. What appears to be the most distinctive property within such systems is the degree of productivity of morphological patterns as defined in this study.

REFERENCES


## APPENDIX

### TEST ITEMS

<table>
<thead>
<tr>
<th>Plural marker</th>
<th>Gender</th>
<th>Test item</th>
<th>Gloss</th>
<th>Prod</th>
<th>Freq Input</th>
<th>Freq CELEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>-s</td>
<td>masc.</td>
<td><em>Clown</em></td>
<td>clown</td>
<td>wp</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td>fem.</td>
<td><em>Pyjama</em></td>
<td>pyjamas</td>
<td>fp</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td>neut.</td>
<td><em>Pizza</em></td>
<td>pizza</td>
<td>p</td>
<td>m</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Oma</em></td>
<td>grandma</td>
<td>fp</td>
<td>h</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Baby</em></td>
<td>baby</td>
<td>fp</td>
<td>m</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Klo</em></td>
<td>toilet</td>
<td>fp</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td>-(e)n</td>
<td>masc.</td>
<td><em>Bub</em></td>
<td>boy</td>
<td>np</td>
<td>m</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td>fem.</td>
<td><em>Hase</em></td>
<td>rabbit</td>
<td>fp</td>
<td>m</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Katze</em></td>
<td>cat</td>
<td>fp</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Uhr</em></td>
<td>watch/clock</td>
<td>p</td>
<td>h</td>
<td>h</td>
</tr>
<tr>
<td></td>
<td>neut.</td>
<td><em>Bett</em></td>
<td>bed</td>
<td>np</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Hemd</em></td>
<td>shirt</td>
<td>np</td>
<td>l</td>
<td>m</td>
</tr>
<tr>
<td>-e</td>
<td>masc.</td>
<td><em>Bus</em></td>
<td>bus</td>
<td>p</td>
<td>m</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Stift</em></td>
<td>pen/pencil</td>
<td>p</td>
<td>m</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Fisch</em></td>
<td>fish</td>
<td>fp</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td>neut.</td>
<td><em>Flugzeug</em></td>
<td>plane</td>
<td>p</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Schaf</em></td>
<td>sheep</td>
<td>wp</td>
<td>m</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Schiff</em></td>
<td>ship</td>
<td>p</td>
<td>l</td>
<td>m</td>
</tr>
<tr>
<td>-e + U</td>
<td>masc.</td>
<td><em>Ball</em></td>
<td>ball</td>
<td>wp</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Hut</em></td>
<td>hat</td>
<td>wp</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Zug</em></td>
<td>train</td>
<td>wp</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td>fem.</td>
<td><em>Bank</em></td>
<td>bench</td>
<td>np</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Maus</em></td>
<td>mouse</td>
<td>np</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Kuh</em></td>
<td>cow</td>
<td>np</td>
<td>l</td>
<td>m</td>
</tr>
<tr>
<td>zero</td>
<td>masc.</td>
<td><em>Pullover</em></td>
<td>pullover</td>
<td>p</td>
<td>m</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Teller</em></td>
<td>plate</td>
<td>p</td>
<td>m</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Tiger</em></td>
<td>tiger</td>
<td>p</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td></td>
<td>neut.</td>
<td><em>Fenster</em></td>
<td>window</td>
<td>p</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Mädchen</em></td>
<td>girl</td>
<td>p</td>
<td>m</td>
<td>h</td>
</tr>
</tbody>
</table>
### Appendix (Cont.)

<table>
<thead>
<tr>
<th>Plural marker</th>
<th>Gender</th>
<th>Test item</th>
<th>Gloss</th>
<th>Prod</th>
<th>Input Freq</th>
<th>CELEX Freq</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Messer</strong></td>
<td></td>
<td>knife</td>
<td>p</td>
<td>l</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>U</strong></td>
<td>masc.</td>
<td>Mantel</td>
<td>coat</td>
<td>np</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td><strong>Hammer</strong></td>
<td></td>
<td></td>
<td>hammer</td>
<td>np</td>
<td>l</td>
<td>1</td>
</tr>
<tr>
<td><strong>Apfel</strong></td>
<td></td>
<td></td>
<td>apple</td>
<td>np</td>
<td>h</td>
<td>l</td>
</tr>
<tr>
<td><strong>Vogel</strong></td>
<td></td>
<td></td>
<td>bird</td>
<td>np</td>
<td>h</td>
<td>m</td>
</tr>
<tr>
<td><strong>Nagel</strong></td>
<td></td>
<td></td>
<td>nail</td>
<td>np</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td><strong>Schnabel</strong></td>
<td></td>
<td></td>
<td>beak</td>
<td>np</td>
<td>l</td>
<td>1</td>
</tr>
<tr>
<td><strong>-er + U</strong></td>
<td>masc.</td>
<td>Schneemann</td>
<td>snowman</td>
<td>np</td>
<td>m</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wurm</td>
<td>worm</td>
<td>np</td>
<td>l</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>neut.</td>
<td>Wald</td>
<td>wood/forest</td>
<td>np</td>
<td>l</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bild</td>
<td>picture</td>
<td>np</td>
<td>m</td>
<td>h</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Haus</td>
<td>house</td>
<td>np</td>
<td>m</td>
<td>h</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Loch</td>
<td>hole</td>
<td>np</td>
<td>h</td>
<td>m</td>
</tr>
</tbody>
</table>

**Key:**

*Productivity:* fp = fully productive; p = productive; wp = weakly productive; np = non-productive.

*Input token frequency:* l = low (<5); m = middle (5–10); h = high (>10).

*CELEX token frequency:* l = low (1–100); m = middle (101–1000); h = high (>1000).