CrossMark

Language and Cognition **8** (2016), 283–313. doi:10.1017/langcog.2014.51
© UK Cognitive Linguistics Association, 2015

Factors considered and ignored in plural acquisition: Frequency rules?*

EUGEN ZARETSKY

Department of Phoniatrics and Pediatric Audiology, Goethe-University Hospital, Frankfurt/Main, Germany, and Institute of Medical Psychology, Goethe-University Hospital, Frankfurt am Main, Germany

BENJAMIN P. LANGE

Department of Media Psychology, Institute for Human-Computer-Media, Faculty of Human Sciences, Julius Maximilian University of Wuerzburg, Germany

HARALD A. EULER

Department of Phoniatrics and Pediatric Audiology, Ruhr-University Hospital, Bochum, Germany

AND

KATRIN NEUMANN

Department of Phoniatrics and Pediatric Audiology, Ruhr-University Hospital, Bochum, Germany

(*Received* 12 *July* 2013 – *Revised* 20 *January* 2014 – *Accepted* 17 *December* 2014 – *First published online* 11 *February* 2015)

ABSTRACT

Several models of regularities behind plural overgeneralization patterns of four-year-old monolingual Germans and bi-/multilingual children acquiring German were empirically examined within the framework of single- and dual-route models. The factors taken into account were frequency, cue validity, productivity, and iconicity of plural allomorphs. In this large-scale cross-sectional study, the results of 893 children tested with a modified, validated version of the speech and language screening *MSS* and of 476 children tested with a validated language test *SETK 3-5* were analyzed. For all overgeneralization models, paired

[*] Address for correspondence: Dr Eugen Zaretsky, Department of Phoniatrics and Pediatric Audiology, Goethe-University Hospital, Frankfurt am Main, Theodor-Stern-Kai 7, House 7A, 60590 Frankfurt am Main, Germany. tel: +49 69 6301 7272; fax: +49 69 6301 5002; mobile: +49 176 8201 9348; e-mail: yevgen.zaretsky@kgu.de

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

groups of expected and not expected answers were compared in the Wilcoxon test. At the early stages of plural acquisition the frequency of plural allomorphs and the simplest phonetic–phonologically based regularities seem to be the crucial factors recognized by children acquiring German. Gender-based plural allomorph frequency was associated with advanced German skills. These findings support the single-route theory with its emphasis on the influence of input characteristics on children's error patterns. No evidence was found for a default status of *-s* or any other plural marker, and hence for the dual-route theory.

KEYWORDS: plural acquisition, German language, single-route models, dual-route models, bilingualism, morphology

1. Introduction

Strategies of plural formation and respective patterns of morpheme OVERGENERALIZATION in first and second language acquisition exemplify a broad range of highly complicated inter-relations between certain parameters. These are FREQUENCY of plural allomorphs in the input language (both types and tokens), their PRODUCTIVITY, CUE VALIDITY, ICONICITY, phonetic–phonologically or morphologically based regularities, analogies with already acquired nouns, or analogy avoidance in favor of default forms, assumed to be used irrespective of any cues like frequency or iconicity of the plural allomorphs.

Previous studies on plural acquisition in German (e.g., Bartke, 1998; Bittner & Köpcke, 2001; Clahsen, 1999; Ewers, 1999; Gawlitzek-Maiwald, 1994) resulted in inconsistent findings about the linguistic features mentioned above. Most of them had insufficiently small sample sizes (e.g., Vollmann, Sedlak, Müller, & Vassilakou, 1997) or used samples comprising children with very broad age groups (Walter, 1975). In some cases, this might have led to unreliable statistical values. For instance, according to Schaner-Wolles (1989), only 20% of the four-year-old Germans in their study were able to produce the correct plural form of the noun *Apfel* 'apple'. However, these 20% referred to only two children out of ten. In one of our databases (4,280 four-year-old monolingual Germans tested with a modified version of the speech and language screening *Marburger Sprachscreening*), more than two-thirds of the children produced this particular plural form correctly (Zaretsky, Euler, Neumann, & Lange, 2014).

Previous research has delivered several hypotheses about the regularities of plural error patterns: (i) Preschool children have not yet internalized rules of plural formation and thus can either reproduce plural forms which they know by heart or avoid plural formation of certain nouns by using synonyms

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

(Mugdan, 1977; Phillips & Lowell, 1980). (ii) The selection of plural allomorphs depends on inter-related factors such as their iconicity, transparency, frequency, productivity, and perceptibility, as well as linguistic knowledge of gender aspects. This hypothesis belongs to the category of the single-route models based on such notions as scheme or analogy learning (Dressler, Mayerthaler, Panagl, & Wurzel, 1987; Ewers, 1999; Korecky-Kröll & Dressler, 2009). (iii) If the plural cannot be formed by means of analogies to already known plural forms, children opt for a default plural form which does not depend on phonetic–phonological features, grammatical characteristics (gender, animacy) or frequency-related aspects of the input language. This hypothesis belongs to the category of the dual-route models (Clahsen, 1999; Clahsen, Rothweiler, Woest, & Marcus, 1992).

In the dual-route models, the plural allomorphs are divided into two groups, namely regular and irregular ones. The *-s* is the only representative of the first group and the other markers belong to the second group. The marker *-(e)n* stands in between, because it is frequently misinterpreted as the default form due to its uniquely high frequency in the input. The only regular plural marker *-s* is believed to be added, both in the language of children and adults, in emergency cases, for instance, when the nouns do not evoke associations with the acquired vocabulary and thus are treated as new material demanding a special marker which does not change the word form by umlauting or accent shifts. Because the vocabulary size and association networks of the preschoolers are rather limited, such emergency cases occur rather often, resulting in a large number of *s*-forms.

According to the dual-route models (Marcus, Brinkmann, Clahsen, Wiese, & Pinker, 1995; Pinker, 1999), the addition of *-s* is governed by a symbolic pluralization rule with clear parallels to the target language. The addition of other plural markers, however, is rather a matter of reproduction of learnt plural forms instead of an application of any internalized rules, or a matter of analogy-based formation. Irregular rules in the dual-route models are, in fact, no rules in a strict sense but analogous reproductions of the acquired plural markers with words having a similar phonetic–phonological structure, the same gender, and other features, whereas the *s*-marker does not depend on any noun characteristics and is supposed to be universally applicable. Dual-route models were originally developed for languages such as English, which has very few exceptions (e.g., *mice*, *men*) to the default forms (e.g., *hats*, *cars*). They were revised for German with its more synthetic structure and with a wide range of plural markers. The model generalization for German resulted, however, in the contra-intuitive finding that the supposed default marker *-s* has, in spite of its high compatibility, a low frequency in types and tokens in the adult language and has, according to numerous reports (e.g., Bittner & Köpcke, 2001; see 'Discussion'), not a dominant, but rather

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

an equivalent position among the plural allomorphs *-(e)n* and *-e* in the overgeneralization patterns.

In the single-route models – Natural Morphology (Dressler et al., 1987; Mayerthaler, 1981; Wurzel, 1984), Cognitive Morphology (Bybee, 1985, 1988; Köpcke, 1993) – the symbolic rules characteristic of the dual-route model are replaced by weighted connections which are constantly adjusted to the regularities extracted from the input and processed by a connectionist pattern associator. Regular and irregular inflections like plural or past participle forms are produced in one uniform way without subdivision into default and irregular forms. The number of features of plural allomorphs taken into account when producing plural forms depends on the degree of language command, with very few at the beginning and a growing range of phonological, semantic, and morphological factors in the advanced stages of language acquisition. Frequently used and easily perceptible plural allomorphs discernible from frequent word-final sound combinations and other morphological elements theoretically applicable in the same position are believed to be acquired earlier and extended in their use to the realm of other plural markers which can be hardly distinguished as such in the input. New cues constantly extracted from the adult language and conflicting priorities of features such as iconicity, productivity, and frequency of plural allomorphs result in constantly changing but systematic overgeneralization patterns.

The proponents of the single-route models reported the overgeneralization patterns which fit into their framework, whereas the proponents of the dual-route models, in spite of similar samples, similar study designs, and items to be tested, observed rather different errors. However, most of these discrepancies are attributable to the small sample sizes, the choice of items, and vague definitions of some terms.

In the present study, several models explaining overgeneralization patterns under consideration of both single- and dual-route theories will be proposed and statistically verified by the comparison of the means of expected and unexpected errors in the answers of the test subjects. Models with significantly higher numbers of expected errors in comparison to the unexpected ones are considered to be effective to a certain extent. The most valid model is the one that provides a significantly higher number of expected answers in comparison with the number of expected answers from all other models.

Our previous studies (Zaretsky, Lange, Euler, & Neumann, 2013; Zaretsky, Neumann, Euler, & Lange, 2013) revealed that neither *-s* nor *-(e)n* can be considered default plural markers. Although this finding will be scrutinized in the 'Results' section again, it is rather to be expected that only the overgeneralization models based on the factors utilized in the single-route theory (plural marker frequency, cue validity, productivity etc.) will reach significant results. Because previous studies (Ewers, 1999; Laaha, Ravid,

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

Korecky-Kröll, Laaha, & Dressler, 2006) delivered no clear clues on which plural allomorph characteristics or their constellations might yield the best results in predicting the overgeneralization patterns, several models need to be examined.

The study design can be summarized as follows. Preschool German children with and without immigration background were given singular forms of various highly frequent real German nouns and some nonce words and were asked to produce plural forms. These plural forms were classified as correct or incorrect uses of the plural allomorphs or as avoidance strategies like repetition of singular forms. Then, the dependence of overgeneralization patterns (that is, wrong uses of plural markers) on various characteristics of German plural allomorphs was analyzed according to eight models stressing various characteristics of these allomorphs. Because monolingual Germans (MO) and bi/-multilingual children (BM) might utilize different overgeneralization patterns, separate analyses were carried out for all children, only for MO, and only for BM. Also, correlations of total scores of correct answers in the language tests with the most relevant plural marker characteristics were calculated because plural tasks were shown to be closely associated with the overall language status of children in one of our previous studies (Zaretsky et al., 2013a).

2. Plural formation in German

Nouns in German are pluralized by means of five suffixes (see Table 1), with or without vowel mutation, also called umlaut. The plural markers *-s* and *-(e)n* are not compatible with umlaut, the plural marker *-er* always requires umlaut. With *-e*, umlaut may or may not occur, which motivated the division of the *e*-marker into *-e* with umlaut and *-e* without umlaut in this study. Some very few nouns are pluralized by means of umlaut alone or by a zero suffix. The choice of the plural allomorph depends on the grammatical gender of the noun, on its phonological form, on morphology (suffix), and on some other factors (Mugdan, 1977). Numerous exceptions must be learned by heart and find their motivation only diachronically.

The simplest and hence the most important phonetic–phonological regularities to be considered in our models correspond to those proposed and statistically backed up by Mugdan (1977): (1) Nouns with schwa as final sound (e.g., *Hose* 'trousers') demand the plural suffix *-(e)n* without exception. (2) With nouns ending in full vowels (e.g., the nonce word *Tulo*), *-s* is very likely. (3) Schwa plus consonant as final sound (e.g., *Apfel* 'apple') requires zero forms or *-n* in most cases. (4) Consonant as final sound (e.g., *Bild* 'picture') requires *-e* more often than any other plural marker. As was demonstrated in Zaretsky et al. (2013b), linguistically more proficient children tended to follow

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

TABLE 1. *Plural morphology of German: examples*

Suffix	Singular	Plural	English translation
-s	*Kino*	*Kinos*	'movie theater'
-e	*Tisch*	*Tische*	'table'
-(e)n	*Eröffnung*	*Eröffnungen*	'opening'
-er	*Brett*	*Bretter*	'board'
-e + umlaut	*Hof*	*Höfe*	'yard'
umlaut	*Laden*	*Läden*	'shop'
no suffix	*Gürtel*	*Gürtel*	'belt'

these regularities, while less proficient children disregarded regularity (4), the most complicated one, and overgeneralized the most frequent German plural allomorph *-(e)n* instead. Regularity (3) refers to the so-called schwa deletion rule which prohibits the use of the schwa in two consecutive syllables; hence overgeneralizations like *Vogeln* 'birds', but not *Vogelen*. According to our test data, this rule, to which no exception exists in adult language, is strictly followed even by the linguistically weakest children and is almost never violated (Zaretsky et al., 2013b).

In order to construct theoretical models which could explain the overgeneralization patterns in our sample, several cues from the language input were examined in a previous study (Zaretsky, Lange, Euler, & Neumann, 2011): general frequency of the plural allomorphs in the adult language, gender-specific frequency (frequency of plural allomorphs with nouns of masculine, feminine, and neuter gender), cue validity (the unambiguousness of an allomorph to signal the plural meaning or, in other words, the relative complexity of being perceived as the plural marker), productivity (the theoretical applicability of a plural allomorph to as many nouns as possible, among others also to nouns which practically require other plural allomorphs), and iconicity (the ability of the plural marker to encode the new meaning – plurality – by a clear-cut new form added to the body of the singular form).

It was demonstrated by Zaretsky et al. (2011) that *(e)n*- and *e*-plurals clearly dominated in the frequency lists (both types and tokens) of the modern German language. For masculine nouns, the suffix *-e* followed by *-e* with umlaut and *-(e)n* dominated among the 1,000 most frequent German noun types. For feminine nouns, the suffixes *-(e)n* followed by *-e* with umlaut and *-s* were the most frequent, and for neuters, the suffixes *-e* followed by *-er* and *-s*. Iconic plural forms (those demanding *-e*, *-er* without umlaut, *-s*, and *-(e)n*) were produced correctly significantly more often than less iconic ones (those demanding *-e* with umlaut, *-er* with umlaut, umlaut alone). As far as cue validity is concerned, maximal values were reached by the plural allomorphs *-s* and *-e*, followed, in this order, by *-(e)n*, *-e* with umlaut, *-er*, and umlaut. One of the concepts of productivity was adapted (Laaha et al., 2006),

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

according to which *-s* is supposed to be productive for any gender and almost any phonetic–phonological environment, *-(e)n* for feminine nouns, *-(e)n* for masculine nouns with schwa as final sound, and *-e* (with and without umlaut) for masculine and neuter nouns. Non-productive are the same plural markers in conditions other than those described above.

Provided statistics (Zaretsky et al., 2011) show that children in our sample considered all the above-mentioned factors and produced significantly more often items with iconic, frequent, productive plural markers as well as markers high on the scale of cue validity. Yet, some contradictions between factors emerged, like high iconicity but low cue validity of the plural marker *-er*, or low frequency but high cue validity of *-s*. In order to examine the weight of single factors, several models will be discussed in the present paper, stressing one or another factor or their constellations.

Wegener (1995) suggested that children might categorize nouns not into three genders but only into feminine vs. non-feminine ones because this dichotomous classification is cognitively less challenging. The motivation for this categorization becomes obvious when comparing the distribution of plural allomorphs according to their gender. For the following calculations, we took the 1,000 most frequently occurring nouns (tokens) from the Corpus of the German language (Leipzig University, Department for Natural Language Processing, 2010). The sum of the differences in the distribution of the raw scores of each plural allomorph (seven allomorphs altogether) among three genders amounts to 299 for the pair masculine vs. neuter nouns, 651 for the pair masculine vs. feminine nouns, and 540 for the pair neuter vs. feminine nouns. It is obvious that the difference between masculine and neuter nouns is the smallest, which could mean that children, at least at an early stage of plural acquisition, do not consistently distinguish between masculine and neuter nouns.

The mean percentage values of the plural allomorphs within the realm of masculine and neuter nouns from the DeReWo list (Institut für Deutsche Sprache, 2009) show the following tendencies: the plural suffix *-e* can be expected to be the most frequent (39.5%), followed by *-e* + umlaut (12.5%), *-s* (10%), *-er* (9.5%), *-(e)n* (7.5%), and umlaut (0%). This finding is of importance for the overgeneralization model 'Feminine vs. non-feminine nouns'; see 'Results'.

3. Methods

3.1. PARTICIPANTS

Participants were children between the ages of 4;0 and 4;11, who were tested in several German cities in the state of Hesse using a new modified, validated version of the language test *Marburger Sprachscreening* (*MSS b*, Euler et al., 2010;

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

Neumann, Holler-Zittlau, van Minnen, Sick, Zaretsky, & Euler, 2011) ($N = 893$) and the validated language test *SETK 3-5* (Grimm, 2001) ($N = 476$). The median age was 4;3 for *MSS b* and 4;2 for *SETK 3-5*.

Each child was classified by a group of speech therapists and linguists as (i) requiring or not requiring educational support in language acquisition, and (ii) requiring or not requiring medical treatment due to speech and/or language impairment. Children who require only educational support are not impaired in the medical sense. Their unsatisfactory command of German is assumed to be mainly related to insufficient or inadequate language input.

Children were classified as BM and MO because BM from the database used here had been shown to score significantly lower than MO in language tests (Zaretsky et al., 2014). Among 893 children tested with *MSS b*, 446 were MO (50%) and 447 BM (50%). Children tested with *SETK 3-5* belonged almost completely to the *MSS b* sample. The largest BM groups in the *MSS b* sample were 78 Turks (9%), 52 Russians (6%), 43 Arabs (5%), and 31 Italians (4%). The classification of the test subjects as MO or BM was based not on the nationality but on the languages spoken by the child and in his or her surroundings. Germans raised bilingually from their birth on were classified as BM.

The age when BM began to learn German was assessed by means of a questionnaire filled out by the daycare center teachers for 188 participants: 65 (35%) had contact to the German language from their birth on, 5 (3%) since their first year of life, 12 (6%) since the second year of life, 22 (12%) since the third year of life, and 84 (45%) since the fourth year of life. Languages spoken at home were documented for 348 BM: 289 (83%) spoke both German and (an)other language(s) at home, 41 (12%) spoke only (an) other language(s), and 18 (5%) spoke only German according to one source (e.g., questionnaire for parents), but some other language(s) according to another source (e.g., questionnaire for daycare center teachers). More information on the sample can be found in Zaretsky et al. (2013b).

3.2. TESTS

MSS b contains subtests on grammar, vocabulary, articulation, phonological short-term memory (repetition of sentences and nonce words), speech comprehension, and fluency. *SETK 3-5* consists of several subtests as well, but only the grammar subtest was of relevance for this study.

The following plural items entered our analysis: (1) items from the *SETK 3-5* test: *der Fisch* 'fish' (singular) – *Fische* (plural), *das Bild* 'picture' – *Bilder*, *der Stuhl* 'chair' – *Stühle*, *das Buch* 'book' – *Bücher*, *die Hand* 'hand' – *Hände*, *das Schiff* 'ship' – *Schiffe*, *das Glas* 'glass' – *Gläser*, *die Gabel* 'fork' – *Gabeln*, *der Vogel* 'bird' – *Vögel*, *der Apfel* 'apple' – *Äpfel*, as well as the nonce words

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

eine Ribane – Ribanen, *ein Tulo – Tulos*, *eine Plarte – Plarten*, *ein Bizo – Bizos*, *eine Tapsel – Tapseln*, *ein Ropf – Röpfe*, *ein Dolling – Dollinge*, *eine Kland – Klände*; (2) the *MSS b* items *Ball* 'ball'– *Bälle*, *Apfel* 'apple' – *Äpfel*, *Auto* 'car' – *Autos*.

The rationale behind the choice of the plural items can be summarized as follows. Nonce words are of special interest for the study because they cannot be retrieved from the mental lexicon and are probably produced by actively used pluralization algorithms based on cues extracted from the input. The *SETK 3-5* provides nonce words which either evoke associations with existing German nouns or correspond to the rules of German phonotactics and thus activate certain pluralization rules or schemata. The *SETK 3-5* items examine all the plural allomorphs except the zero allomorph. Because the zero allomorph cannot be distinguished from the repeated singular forms without context, such items were not included in the analysis. Because not all children could cope with the nonce words, and due to the fact that low compliance in this *SETK 3-5* subtest occurred rather frequently in the pilot test sessions, real words were also integrated into the test battery.

Children were tested in the daycare centers after their parents had signed informed consent on the study participation. For the elicitation of plural forms, children were shown pictures with one or several objects and were asked questions like "Look, here is one chair and there we can see three …?" All in all, 3,282 overgeneralizations and zero forms in the *SETK 3-5* and 852 such forms in the *MSS b* were documented. Of interest were also *MSS b* tasks on the accusative and dative forms in which the command of the grammatical gender can be assessed. Answers containing articles and prepositions (e.g., *aus dem Sandkasten* 'from the sandbox') were elicited in this case. The child was shown certain situations on a large, colorful picture depicting playing children and was asked questions on these situations.

3.3. STATISTICAL ANALYSES

All wrong children's responses in the tests entered the analyses as error patterns (*Apfels*, *Äpfels, Apfeler*, *Äpfeler*, etc. 'apples'), including zero form (*Apfel*), other words (*Korb* 'basket'), deformed forms (*Applemle*), and quantifying responses like 'many', 'three', 'some' without plural forms. The categorization of the error patterns was undertaken by one research assistant and the first author independently from one another. Very good inter-rater reliability was given (Cohen's $\kappa \geq .94$ for all data entries in *MSS b* and *SETK 3-5*).

Normal distribution was checked for using the Kolmogorov–Smirnov test in order to decide whether parametric or non-parametric analyses on group differences would have to be conducted. Because the data were not normally distributed (ps $< .05$), non-parametric tests were used: the Wilcoxon test for

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

comparisons of paired groups of overgeneralization models, the Mann–Whitney *U* test for comparisons of independent groups, and Spearman correlations to assess associations between plural marker characteristics and total scores of correct answers, zero forms, and actively used plural allomorphs. Point-biserial correlations were utilized to examine the association between the most relevant plural marker characteristics and the classification of all children as needing educational and/or medical help in learning/acquiring German. Sociolinguistic characteristics of children who did not produce any correct or wrong plural forms were compared with characteristics of all other children by means of the chi-square test.

All data were processed using SPSS 15 (International Business Machines Corp., New York, USA): $p < .05$ was considered to indicate a statistically significant difference.

3.4. OVERGENERALIZATION MODELS

All the proposed overgeneralization models are based on the same rules:

1. Wrong answers were classified into the following categories: overgeneralizations of *-e*, *-e* plus umlaut, umlaut, *-(e)n*, *-s*, *-er*, and zero forms.
2. Zero forms were not considered to be overgeneralizations of the zero plural allomorph (-Ø), but rather an avoidance strategy (cf. Zaretsky et al., 2013b, for argumentation).
3. Only wrong forms of plural items were considered, because in the case of correct forms it could not be ruled out that memorized forms were merely reproduced.
4. Due to (3), predictions for structurally similar plural items may vary. For instance, in the first model ('Gender only'), for the masculine noun *Fisch* 'fish' the plural form *Fischen* was predicted because the most probable form *Fische* corresponds to the correct form (see the gender-based plural marker distribution in Section 2, 'Plural formation in German'), whereas the prediction for the masculine noun *Stuhl* 'chair' was *Stuhle* because masculine nouns occur with *-e* more often than with any other plural marker and the correct form of *Stuhl* is *Stühle* (not *-e*, but *-e* with umlaut).
5. Each model proposed only one appropriate plural allomorph which is most likely to be overgeneralized with a certain item.

Only the frequency of plural allomorphs in the input unambiguously predicts the respective pattern of overgeneralization, that is, one can expect the most frequent suffix to be followed by the second most frequent suffix, and so forth. In contrast, iconicity, productivity, and cue validity are comparatively vague

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

indicators of overgeneralization patterns. However, one would expect that any model which is supposed to explain overgeneralization patterns accounts for these factors as well.

4. Results

Several potential models for selecting plural allomorphs were compared with each other based on the data obtained from *SETK 3-5*. Factors taken into account were those verified statistically by Zaretsky et al. (2011) and described in Section 2 'Plural formation in German': frequency (general vs. gender-specific frequency), cue validity, productivity, iconicity of plural allomorphs, as well as the most basic phonetic–phonological regularities.

The first model ('Gender only') considers gender-based plural allomorph distribution as the only factor and ignores all phonetic–phonologically based regularities. According to this model, the following overgeneralizations were expected for the real nouns: (1) masculine nouns: *Fischen*, *Stuhle*, *Vogele*, *Apfele*; (2) neuter nouns: *Bilde*, *Buche*, *Schiffer*, *Glase*; (3) feminine nouns: *Handen*, *Gäbele*. Some forms (e.g., *Apfele*, *Vogele*, *Gäbele*) contradict the schwa deletion rule, iconicity (*Gäbele*) or the principle of cue validity (*Schiffer*). Results of all models are given later in Table 3.

Apart from gender aspects, the second model ('Gender + schwa') takes the schwa deletion rule into consideration, which is the only phonetic–phonological rule strictly followed by our sample (Zaretsky et al., 2013b). The 'Gender + schwa' model predicts the following overgeneralizations for the *SETK 3-5* real nouns: (1) masculine nouns: *-(e)n* for *Vogel*, *Apfel*, *Fisch*, *-e* for *Stuhl*; (2) neuter nouns: *-e* for *Buch*, *Glas*, *Bild*, *-er* for *Schiff* (violation of the principle of cue validity); (3) feminine nouns: *-(e)n* for *Hand*, *-s* for *Gabel*.

In the model 'Feminine vs. non-feminine nouns', we followed Wegener's (1995) assumption that children do not categorize nouns into three genders but only into feminine vs. non-feminine ones. The motivation for this division and percentage values for the distribution of plural allomorphs in the feminine and non-feminine groups were given in Section 2 'Plural formation in German'. Taking these percentage values into account and considering the schwa deletion rule, we would expect the following overgeneralization patterns in the model 'Feminine vs. non-feminine nouns': *Fisches*, *Bilde*, *Stuhle*, *Buche*, *Schiffs*, *Glase*, *Vogels*, *Apfels*. The expected suffixes for feminine nouns are the same as in the 'Gender + schwa' model. The principles of cue validity and iconicity are not violated.

The next model neglects gender-based rules for the distribution of plural allomorphs ('Gender neglect' model). Instead, the following rules, which are based on the simple phonetic–phonological sound patterns of the nouns, are taken into consideration: (1) *-e* and *-(e)n* as the most frequent and the earliest

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

plural allomorphs (Bittner & Köpcke, 2001) are only seldom substituted by other allomorphs, because they are constantly enforced by the input the children receive from the adult language. (2) Due to the high frequency of *-e* and the high applicability of *-s* in the input, their iconicity, and cue validity, *-e* + umlaut is not substituted by other suffixes but by *-e* without umlaut (Kauschke, Kurth, & Domahs, 2011). The plural allomorph *-(e)n* is substituted by *-s*, the only plural allomorph which is spreading in the linguistically more proficient groups (Zaretsky et al., 2013b). (3) The schwa deletion rule, which requires a non-syllabic plural marker after a schwa sound (*-s* or *-n*), is never violated (cf. Kauschke et al., 2011). (4) In nouns ending in a full vowel, only *-s* is added, because this is one of the easiest plural rules (Wegener, 1995). (5) In nouns ending in *-e*, only *-n* is added, because this is the easiest plural suffixation rule in German (Clahsen et al., 1992; cf. Kauschke et al., 2011). (6) As the umlaut alone is minimally iconic, not productive, rare in the input, and low on cue validity, it does not suffice for marking plural. Hence, overgeneralizations of more iconic plural allomorphs can be expected (Ewers, 1999; Kauschke et al., 2011; Laaha et al., 2006; Schaner-Wolles, 1989; Szagun, 2001). (7) In cases where principles (2) and (4) do not apply, the plural allomorph which is most frequent in the input language will be overgeneralized, namely *-(e)n*. (8) As in all previous studies, the most common wrong answer will be the zero form (Clahsen et al., 1992; Schaner-Wolles, 2001; Szagun, 2001).

Our eight rules in the 'Gender neglect' model take Mugdan's (1977) principles (1), (2), and (3) into consideration and do not account for (4) (see Section 2 'Plural formation in German'). That is, the model predicts nouns with a full vowel as a word final sound to take *-s*, nouns with a schwa in the word final position to take *-n*, nouns with the schwa in the last syllable before a consonant to take some non-syllabic plural suffix. The model, however, does not predict nouns ending in a consonant to take *-e*. Although the suffix *-e* did occur very often with items like *Fisch* and *Bild*, it did not dominate in the error patterns: (1) BM: *M*s = 0.6 *e*-overgeneralizations, *SD* = 0.91 vs. 1.4 other overgeneralizations, *SD* = 2.02, $N = 176$, $Z = -4.50$, $p < .001$, (2) MO: *M*s = 0.6, *SD* = 0.77 vs. 1.3, *SD* = 1.50, $N = 300$, $Z = -7.14$, $p < .001$.

The correctness of principles (1) to (3) by Mugdan (1977) is controlled by the data presented in Table 2. It sums up our major findings on the applicability of the plural allomorphs with *SETK 3-5* items.

Our eight rules (see above) led us to expect the following overgeneralizations in the 'Gender neglect' model. The item *Gabel* would mostly be overgeneralized by adding *-s*, because the schwa sound only allows *-s* and *-n* as plural suffixes, and *-n* is the correct answer anyway. *Vogel* and *Apfel* would get the suffix *-(e)n* (in the form *-n*), which is the most frequent suffix in the children's input and does not contradict the schwa deletion rule. *Hand* and *Stuhl* would get

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

TABLE 2. SETK 3-5 *with four-year-old children. Error patterns which explain 90% of all errors. 1 = most frequent errors, 4 = least frequent errors. The hyphen (-) marks the correct plural form.*

	-er	**-Ø**	**-(e)n**	**-s**	**-e**	**e + uml**	**uml**
Fisch 'fish'		1	2		-		
Bild 'picture'	-	1	2				
Stuhl 'chair'		1	3		2	-	
Buch 'book'	-	1	2		3		
Hand 'hand'		1	3		2	-	
Schiff 'ship'		1	2		-		
Glas 'glass'	-	1	3		2		
Gabel 'fork'		1	-	2			
Vogel 'bird'		1	2	3			-
Apfel 'apple'		1	2	3			-
Nonce words:							
Ribane		1	-				
Tulo		1		-			
Plarte		1	-				
Biwo		1		-			
Tapsel		1	-	2			
Ropf		1	4	3	2	-	
Dolling		1		2	-		
Kland		1	4	3	2	-	

the second commonest suffix *-e* without umlaut because umlaut as a minimally iconic part of plural markers is hardly perceived as such and thus is frequently left out. *Bild*, *Fisch*, *Schiff*, *Buch*, and *Glas* would get the suffix *-(e)n*, which is the most frequent in the input and does not contradict any other point. Only the real nouns are considered to assure comparability with the previous three models. Results are presented in Table 3.

A comparison between the 'Feminine vs. non-feminine nouns' model and the 'Gender neglect' model is also possible for the nonce words in *SETK 3-5*, which were presented with indefinite articles (they do not allow one to distinguish masculine and neuter nouns but allow the distinction of feminine and non-feminine ones in German). In the 'Gender neglect' model, one would expect the overgeneralizations *Ribanes*, *Tulo*(*e*)*n*, *Plartes*, *Biwo*(*e*)*n*, *Tapsels*, *Ropfe*, *Dollings*, *Klande*. In the model 'Feminine vs. non-feminine nouns' one would expect *Tuloe*, *Ribanes*, *Biwoe*, *Plartes*, *Tapsels*, *Ropfe*, *Dollings*, *Klande*. In the case of the item *Dolling*, it was assumed for both models that the schwa deletion rule is mistakenly followed by the children. We summed up the expected and non-expected answers for both nonce and real nouns (Table 3). The two gender-based models ('Gender only' and 'Gender + schwa') cannot be examined on the basis of nonce words because of the indefinite article forms.

Furthermore, several other models were tested for all *SETK 3-5* items. The fifth model ('Feminine vs. non-feminine nouns + schwa + *-(e)n*') accounted

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

TABLE 3. *Hypothetical models for overgeneralization patterns in* SETK 3-5 *compared by mean values and Wilcoxon test.* N = *457 (295 monolingual Germans = MO, 162 bi-/multilingual children = BM).*

		Real words from *SETK 3-5*				All *SETK 3-5* items (real and nonce words)					
		Gender only	Gender + schwa	Feminine vs. non-feminine nouns	Gender neglect	Feminine vs. non-feminine nouns	Gender neglect	Feminine vs. non-feminine nouns + schwa + -(e)n	Gender neglect + -s	Gender neglect + schwa	Gender neglect + -e
Means (expected answers)/SD	All	0.3/0.63	0.6/0.91	0.5/0.86	0.6 /1.16	1.0/1.28	1.2/1.45	1.0/1.14	0.9/1.17	1.2/1.40	1.2/1.32
	MO	0.2/0.54	0.5/0.91	0.5/0.91	0.5/0.87	1.1/1.36	1.1/1.26	1.0/1.09	1.1/1.27	1.2/1.38	1.1/1.20
	BM	0.5/0.74	0.8/1.07	0.5/0.77	0.9/1.51	0.9/1.10	1.3/1.72	1.0/1.22	0.7/0.94	1.2/1.45	1.3/1.51
Means (not expected answers)/SD	All	0.7/1.25	0.5/0.97	0.5/1.13	0.4/0.78	0.9/1.42	0.8/1.06	1.0/1.45	1.1/1.57	0.9/1.08	0.8/1.89
	MO	0.6/1.04	0.4/0.81	0.4/0.82	0.4/0.77	0.8/1.15	0.8/1.04	0.9/1.27	0.9/1.25	0.8/0.96	0.8/1.13
	BM	0.9/1.55	0.7/1.20	0.9/1.49	0.5/0.79	1.2/1.77	0.8/1.10	1.2/1.72	1.5/1.96	1.0/1.25	0.9/1.30
Z	All	–7.55***	–2.31*	–0.07	–3.53***	–1.89	–4.81***	–0.07	–0.66	–4.37***	–4.85***
	MO	–6.92***	–1.89	–1.75	–1.35	–3.56***	–3.52***	–1.12	–2.60**	–4.40***	–3.24**
	BM	–3.61***	–1.35	–2.20*	–3.69***	–1.77	–3.36**	–1.28	–4.69***	–1.25	–3.80***

NOTE: * $p < .05$; ** $p < .01$; *** $p < .001$.

for the 'feminine vs. non-feminine'-based regularity requiring *-e* with non-feminine nouns, also under consideration of the schwa deletion rule and the default-like status of *-(e)n*, hence coming close to the 'Gender neglect' model. The expected forms were *Fische*, *Bilde*, *Stuhle*, *Buche*, *Hande*, *Schiffe*, *Glase*, *Gabels*, *Vogeln*, *Apfeln*, *Ribanes*, *Tulon*, *Plartes*, *Biwon*, *Tapsels*, *Ropfe*, *Dollinge*, *Klanden*.

The sixth model, 'Gender neglect + -s', is similar to 'Gender neglect', but replaces *-(e)n* with *-s* as the default-like plural: *Fisches*, *Bildes*, *Stuhle*, *Buchs*, *Hande*, *Schiffs*, *Glases*, *Gabels*, *Vogels*, *Apfels*, *Ribanes*, *Tulon*, *Plartes*, *Biwon*, *Tapsels*, *Ropfe*, *Dollings*, *Klande*.

The seventh model, 'Gender neglect + schwa', corresponds basically to 'Gender neglect', but with a stronger consideration of the schwa deletion rule. 'Gender neglect + schwa' took as a starting point the assumption that the transition from the preference for *-(e)n* to the preference for *-s* begins with those items which unambiguously forbid the syllabic suffixes, that is, items governed by the schwa deletion rule. Therefore, one would expect the error patterns *Gabels*, *Vogels*, *Apfels*, *Tapsels*, and *Dollings* (if we assume that the vowel in the last syllable is regularly mistaken for schwa in *Dolling*), but *Fischen*, *Bilden*, *Buchen*, *Schiffen*, *Glasen*, *Tulon*, *Biwon*. The assumption that children tend to overgeneralize *-s* with items that have a schwa in the last syllable is based on the observation that our *SETK 3-5* sample overgeneralized *-s* significantly more frequently with items governed by the schwa deletion rule than with other *SETK 3-5* items: 0.05 *s*-overgeneralizations with the items demanding schwa deletion (means of overgeneralizations divided by the total number of the items) ($SD = 0.06$ vs. 0.02 with other items, $SD = 0.02$, $N = 458$, $Z = -4.99$, $p < .001$).

The last model, 'Gender neglect + -e', basically corresponds to the 'Gender neglect' model, but takes into account that there is a tendency to overgeneralize *-e* with those items which contain the vowels *o*, *a*, *u*, that is, vowels which can potentially be umlauted (cf. Table 2). This results in two discrepancies from the 'Gender neglect' model: overgeneralizations *Glase* instead of *Glasen* and *Buche* instead of *Buchen*.

In Table 3, the means of expected and not expected overgeneralizations per child in all models are compared.

Next, the means of the expected errors according to different models were compared (Table 4). The model which dominates over other models for this criterion was supposed to provide a better explanation for the regularities behind the overgeneralization patterns.

Further, the influence of the plural allomorph features discussed above (and typically related to the single-route models) on the choice of plural markers in *SETK 3-5* was tested. All the plural allomorphs were classified as belonging or not belonging to the groups of iconic, productive, frequent (general and

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

TABLE 4. *Means of the expected errors in various explanatory models in* SETK 3-5 *compared with Wilcoxon test. N = 457 (295 monolingual Germans = MO, 162 bi-/multilingual children = BM).*

	Means			SD			Z		
	All	MO	BM	All	MO	BM	All	MO	BM
All *SETK 3-5* items									
M3 vs. M4	1.1 vs. 1.2	1.2 vs. 1.1	0.9 vs. 1.3	1.28 vs. 1.45	1.36 vs. 1.26	1.10 vs. 1.72	–1.86	–0.63	–3.40**
M3 vs. M5	1.1 vs. 1.0	1.2 vs. 1.0	0.9 vs. 1.0	1.28 vs. 1.14	1.36 vs. 1.09	1.10 vs. 1.22	–0.96	–2.56*	–1.83
M3 vs. M6	1.1 vs. 1.0	1.2 vs. 1.1	0.9 vs. 0.7	1.28 vs. 1.18	1.36 vs. 1.27	1.10 vs. 0.94	–3.90***	–2.43*	–3.07**
M3 vs. M7	1.1 vs. 1.2	1.2 vs. 1.2	0.9 vs. 1.2	1.28 vs. 1.40	1.36 vs. 1.38	1.10 vs. 1.45	–2.36*	–0.38	–2.78**
M3 vs. M8	1.1 vs. 1.2	1.2 vs. 1.1	0.9 vs. 1.3	1.28 vs. 1.33	1.36 vs. 1.20	1.10 vs. 1.52	–2.35*	–0.76	–4.48***
M4 vs. M5	1.2 vs. 1.0	1.1 vs. 1.0	1.3 vs. 1.0	1.45 vs. 1.14	1.26 vs. 1.09	1.72 vs. 1.22	–4.21***	–2.60**	–3.43**
M4 vs. M6	1.2 vs. 0.9	1.1 vs. 1.1	1.3 vs. 0.7	1.45 vs. 1.18	1.26 vs. 1.27	1.72 vs. 0.94	–4.37***	–0.40	–5.48***
M4 vs. M7	1.2 vs. 1.2	1.1 vs. 1.2	1.3 vs. 1.2	1.45 vs. 1.34	1.26 vs. 1.38	1.72 vs. 1.45	–0.91	–1.23	–3.34**
M4 vs. M8	1.2 vs. 1.2	1.1 vs. 1.1	1.3 vs. 1.3	1.45 vs. 1.33	1.26 vs. 1.20	1.72 vs. 1.51	–0.90	–0.75	–0.55
M5 vs. M6	1.0 vs. 0.9	1.0 vs. 1.1	1.0 vs. 0.7	1.14 vs. 1.18	1.09 vs. 1.27	1.22 vs. 0.94	–1.46	–1.38	–4.17***
M5 vs. M7	1.0 vs. 1.2	1.0 vs. 1.2	1.0 vs. 1.2	1.14 vs. 1.40	1.09 vs. 1.38	1.22 vs. 1.45	–3.27**	–2.90**	–1.71
M5 vs. M8	1.0 vs. 1.2	1.0 vs. 1.1	1.0 vs. 1.3	1.14 vs. 1.33	1.09 vs. 1.20	1.22 vs. 1.51	–5.51***	–2.79**	–4.89***
M6 vs. M7	0.9 vs. 1.2	1.1 vs. 1.2	0.7 vs. 1.2	1.17 vs. 1.40	1.27 vs. 1.38	0.94 vs. 1.45	–6.17***	–3.48***	–5.03***
M6 vs. M8	0.9 vs. 1.2	1.1 vs. 1.1	0.7 vs. 1.3	1.17 vs. 1.33	1.27 vs. 1.20	0.94 vs. 1.51	–4.48***	–0.38	–6.09***
M7 vs. M8	1.2 vs. 1.2	1.2 vs. 1.1	1.2 vs. 1.3	1.40 vs. 1.33	1.38 vs. 1.20	1.45 vs. 1.51	–0.29	–1.33	–2.41*
SETK 3-5 real words only									
M1 vs. M2	0.3 vs. 0.6	0.2 vs. 0.5	0.5 vs. 0.8	0.63 vs. 0.91	0.54 vs. 0.80	0.74 vs. 1.07	–9.17***	–7.5***	–5.3***
M1 vs. M3	0.3 vs. 0.5	0.2 vs. 0.5	0.5 vs. 0.5	0.63 vs. 0.86	0.54 vs. 0.91	0.74 vs. 0.77	–5.99***	–6.22***	–0.72
M1 vs. M4	0.3 vs. 0.6	0.2 vs. 0.5	0.5 vs. 0.9	0.63 vs. 1.16	0.54 vs. 0.87	0.74 vs. 1.51	–6.60***	–5.21***	–4.23***
M2 vs. M3	0.6 vs. 0.5	0.5 vs. 0.5	0.8 vs. 0.5	0.91 vs. 0.86	0.87 vs. 0.80	1.07 vs. 0.77	–2.33*	–0.14	–4.00***
M2 vs. M4	0.6 vs. 0.6	0.5 vs. 0.5	0.8 vs. 0.9	0.91 vs. 1.16	0.80 vs. 0.87	1.07 vs. 1.51	–1.40	–0.53	–2.20
M3 vs. M4	0.5 vs. 0.6	0.5 vs. 0.5	0.5 vs. 0.9	0.86 vs. 1.16	0.91 vs. 0.87	0.77 vs. 1.51	–2.06*	–0.35	–3.39**

NOTES: * $p < .05$; ** $p < .01$; *** $p < .001$.

M1 – Gender only, M2 – Gender + schwa, M3 – Feminine vs. non-feminine nouns, M4 – Gender neglect, M5 – Feminine vs. non-feminine nouns + schwa + -(e)n, M6 – Gender neglect + -s, M7 – Gender neglect + schwa, M8 – Gender neglect + -e.

gender-specific frequency), and high on cue validity allomorphs for each *SETK 3-5* item, with equal numbers of allomorphs in each group, thus resulting in five classifications (new variables) per item and five corresponding sum variables. Each pair of such sum variables was compared by the Wilcoxon test for MO and BM groups. In the answers of MO, the gender (feminine vs. non-feminine) based frequency of plural allomorphs was preferred to iconicity, productivity, general frequency, and cue validity (all ps < .001, Ns = 295). BM preferred general frequency of plural allomorphs over iconicity, productivity, and cue validity (ps < .05, N = 162). The preference over gender (feminine vs. non-feminine) based frequency was numerically higher without being statistically significant. Thus, in both groups, the frequency-related features of plural allomorphs dominated over iconicity, productivity, and cue validity. Low difficulty levels of some *SETK 3-5* items (*Fisch*: 75% correct answers, *Bild*: 68%, *Stuhl*: 70%) did not allow the exclusion of correct answers in such calculations.

Next, mean values of the total scores of all plural allomorph characteristics were calculated for each *SETK 3-5* participant, that is, how many times the child, for instance, produced plural forms which do not contradict the principles of cue validity in the total of 18 items. Mean values of the general and gender-based frequency per child were, according to the Wilcoxon test, significantly higher than the mean values of cue validity, productivity, and iconicity both for the whole sample (Z = –12.26, p < .001, N = 457), for the MO group (Z = –12.20, p < .001, N = 290), and for the BM group (Z = –3.59, p < .001, N = 167). According to the same test, mean values of both frequencies yielded significantly higher results in comparison with other plural allomorph characteristics both for MO and BM groups, but only in the MO group the mean values of the gender-based frequency were significantly higher than the mean values of all other characteristics including general frequency (Table 5).

High relevance of the gender-based frequency for overgeneralizations in the MO group might mean that the command of this characteristic is more closely associated with the total number of correct answers in *SETK 3-5* and *MSS b* than in the case of the general frequency. Also, one would expect that children with a good command of the gender-based frequency use more plural allomorphs and less zero forms. Indeed, as is shown in Tables 6, 7, and 8, our data support these assumptions both for all children and for the subgroups.

A Spearman correlation between the general frequency in *SETK 3-5* and the number of correct forms in *MSS b* was lower than the correlation between gender-based frequency in *SETK 3-5* and the same *MSS b* tasks (r_s = 0.527 vs. .668, ps < .001, Ns = 298). A point-biserial correlation between gender-based frequency and the classification of all children as needing educational or medical help in acquiring/learning German was higher for the gender-based frequency than for the general frequency: (1) children needing educational

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

TABLE 5. *Mean values of plural allomorph characteristics per child compared by the Wilcoxon test for paired groups: frequency vs. cue validity, productivity, and iconicity.* N = *357 (290 monolingual Germans = MO, 167 bi-/multilingual children = BM).*

		Gender-based frequency vs. cue validity, productivity, iconicity	Gender-based frequency vs. cue validity, productivity, iconicity, general frequency	General frequency vs. cue validity, productivity, iconicity
Z	All	–12.05***	–11.35***	–5.96***
	MO	–13.04***	–12.72***	–3.76***
	BM	–2.17*	–1.51	–5.17***

NOTE: * $p < .05$; ** $p < .01$; *** $p < .001$.

support (r_{pb} = –0.614 vs. –0.495); (2) children needing medical help (r_{pb} = –0.228 vs. –0.149); (3) children needing both educational support and medical help (r_{pb} = –0.598 vs. –0.461, ps ≤ .05, Ns = 167).

Not all children produced plural forms. In *MSS b*, 95 out of 893 children repeated singular forms in all three cases (11%). These children belonged predominantly to the BM group ($\chi^2(1, N = 882) = 73.77$, $p < .001$) and, within the BM group, they were mostly Arabs ($\chi^2(1, N = 453) = 5.02$, $p = .025$) and, marginally significantly, Turks ($\chi^2(1, N = 453) = 3.16$, $p = .075$). Children who repeated singular forms in all three cases scored significantly lower in *MSS b* than other children according to the Mann–Whitney U test (Z = –9.69, $p < .001$, N = 77 vs. 516). Although they did not attend daycare centers for a significantly shorter period of time than other children ($p > .05$), their German skills were estimated by daycare center teachers significantly lower than those of other children both at the beginning of their daycare center attendance (Z = –3.16, $p = .002$, N = 47 vs. 157) and when the language tests were conducted (Z = –3.40, $p = .001$, N = 48 vs. 161). They were also significantly more often classified as needing educational and/or medical help in acquiring/learning German in comparison with children who produced at least one correct or wrong plural form ($\chi^2(1, N = 892) = 96.71$, $p < .001$).

In the 'Gender neglect' model, rules (1), (2), and (6) suggest that we attribute certain characteristics of the default plural marker to the plural allomorphs *-(e)n* and *-s*, because both are overgeneralized if other suffixes do not match. Indeed, the wide scope of use of these suffixes induces us to assume that they are universally applicable. In *SETK 3-5*, for instance, for 8 of 14 items, more than 50% of all overgeneralizations were those with *-(e)n*. For 5 of 16 items, more than 50% of all overgeneralizations were those with *-s*. Comparing BM and MO shows that *-s* possesses some characteristics of a default form for MO, while *-(e)n* has, with limitations, some characteristics of a default form for BM. This can be concluded, among other things, from

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

TABLE 6. *Spearman correlations between plural allomorph frequency and total scores of* SETK 3-5. N = *455 (288 monolingual Germans = MO, 167 bi-/multilingual children = BM). All ps < .001.*

	General frequency	Gender-based frequency
All	.692	.877
MO	.655	.782
BM	.798	.938

the fact that MO overgeneralized *-s* significantly more often than the second most common plural allomorph *-e* (Mann–Whitney *U* Test: 0.6 occurrences of *-e*, $SD = 0.78$ vs. 0.8 occurrences of *-s*, $SD = 1.10$, $N = 295$, $p < .01$, $Z = -2.87$), whereas BM overgeneralized *-(e)n* significantly more often than the second most common plural marker *-e* (1.7 occurrences of *-e*, $SD = 1.63$ vs. 2.5 occurrences of *-(e)n*, $SD = 2.82$, $N = 162$, $p < .05$, $Z = -2.23$). However, both *-(e)n* and *-s* were hardly overgeneralized for certain items (e.g., *Dolling* with *-(e)n*, *Bücher* with *-s*).

5. Discussion

Previous studies on the regularities in plural acquisition in German resulted in contradictory findings concerning overgeneralization patterns and the most frequent plural markers. These contradictions might be traced back to, among other things, small sample sizes and a suboptimal choice of test items. The present study tried to close this research gap by utilizing a large sample including both children with and without immigration background and by choosing test items which were supposed to elicit all identifiable plural markers (zero forms cannot be distinguished from repeated singular forms and thus were not a subject of this study). To sum up the study results with one sentence, the overgeneralization patterns of most children were based on the simplest phonetic–phonological regularities ('Gender-neglect' models) and on the frequency of plural markers in the input ('Feminine vs. non-feminine nouns' model) and thus supported the single-route theories rather than the dual-route ones.

Several factors influencing plural acquisition and formation regularities have been examined in this paper, namely frequency of plural allomorphs, gender aspects, phonetic–phonological patterns as well as iconicity, cue validity, and productivity of plural allomorphs. Other parameters (e.g., morphological patterns) were not examined due to the characteristics of the test items. For bi-/multilingual children (BM), the non-consideration of gender aspects in favor of the easiest rules of plural formation based on the phonetic–phonological sound patterns of a word was the best explanation for the overgeneralization patterns. Monolingual German children (MO) considered both gender aspects

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

TABLE 7. *Spearman correlations between the plural allomorph frequency and the total scores of all* MSS b *subtests.* N = *298 (131 monolingual Germans = MO, 167 bi-/multilingual children = BM).*

	Speech comprehension	Vocabulary	Grammar	Only plural	Word repetition	Sentence repetition	Articulation	Total score
General frequency								
All	.401***	.547***	.579***	.478***	.292***	.468***	.175**	.581***
MO	.314***	.459***	.464***	.256**	.321***	.442***	.145	.522***
BM	.388***	.584***	.647***	.568***	.260**	.459***	.141	.601***
Gender-based frequency								
All	.445***	.670***	.721***	.659***	.352***	.589***	.246***	.728***
MO	.315***	.538***	.531***	.375***	.453***	.472***	.135	.594***
BM	.424***	.663***	.749***	.716***	.283***	.583***	.195*	.722***

NOTE: * $p < .05$; ** $p < .01$; *** $p < .001$.

TABLE 8. *Spearman correlations between the plural allomorph frequency and total numbers of plural allomorphs and zero forms in* SETK 3-5. N = *457 (290 monolingual Germans = MO, 167 bi-/multilingual children = BM). All ps < .001.*

	Zero forms	Plural allomorphs (types)	Plural allomorphs (tokens)
General frequency			
All	–0.486	0.687	0.822
MO	–0.517	0.574	0.751
BM	–0.440	0.777	0.925
Gender-based frequency			
All	–0.569	0.816	0.943
MO	–0.686	0.734	0.918
BM	–0.453	0.851	0.955

(or, rather, the more obvious and simply perceptible division into feminine and non-feminine nouns) and the same simple phonetic–phonologically based probabilistic patterns which were extracted from the input.

According to Table 3, it remains uncertain which of the models 'Feminine vs. non-feminine nouns', 'Gender neglect', and 'Gender neglect + schwa' can be called the most appropriate one for the MO group, because in all three cases the differences between the expected and non-expected error patterns were highly significant. For the BM sample, the model 'Gender neglect + -e' yields the most appropriate explanation, followed by 'Gender neglect'. In Table 4, the Wilcoxon test revealed that for MO, the models 'Feminine vs. non-feminine nouns' and 'Gender neglect + schwa' yielded significantly better predictions in two out of four cases and thus provided better explanations than other models, but without clear dominance. For BM, 'Gender neglect', 'Gender neglect + schwa', and 'Gender neglect + -e' provided the best solution, with 'Gender neglect' and 'Gender neglect + -e' clearly dominating. For the entire sample, the same three models prevailed, but none of them dominated clearly. Consequently, the models followed by the BM and MO groups differ, although both models 'Gender neglect + schwa' and 'Gender neglect + -e' are modifications of the model 'Gender neglect'.

Most of the models were based on the premises of the single-route theory. Exceptions were the models 'Gender neglect + -s' and 'Feminine vs. non-feminine nouns + schwa + -(e)n', with either *-s* or *-(e)n* as default plural forms. Both of them failed to account for the regularities behind the overgeneralization patterns in our sample. It should be noted, however, that the phonotactics of the chosen test items (nonce words) probably did not always provoke the use of emergency plural forms because of more or less clear associations with other

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

existing German words. On the other hand, the assumption that associations with real German words must block the use of emergency plural forms can be falsified by the fact that even the most obvious associations like *Kland – Hand* 'hand' in *SETK 3-5* were not recognized by the children in our sample. In case of the item *Kland*, it resulted in 96% of wrong (non-expected) plural forms.

The sixth model, 'Gender neglect + -s', paid tribute not only to the dual-route theories but also to the idea of Stemberger (1999) that Germans generally tend to overgeneralize *-s* not due to its default status, but rather as a result of its monosyllabicity. Because monosyllabic words dominate in German vocabulary and *-s* does not add a new syllable to the plural form, Germans are supposed to prefer *s*-overgeneralizations. Stemberger's assumption does not explain, however, the almost total lack of monosyllabic plural forms in German. Furthermore, Stemberger (1999) stresses the influence of foreign languages like English, French, and Spanish, which use *-s* for pluralization and are widely taught at German schools, on the pluralization of nonce words, obviously perceived as foreign ones. Although such influence cannot be excluded (on the contrary, one could add that these foreign languages are spoken more or less widely in the German daycare centers as well), the model 'Gender neglect + -s' could not enter the group of the most important ones, and, as we have demonstrated in another recent study (Zaretsky, Hey, & Lange, unpublished observations), there is no evidence that even German adults, in spite of their often good command of English, utilize *-s* with nonce words as a default plural marker. Rather, they overgeneralize the same three plural markers which can be found in the answers of MO and BM: *-(e)n*, *-s*, and *-e*. The items in the study by Zaretsky et al. (unpublished observations) were taken from studies by the proponents of the dual-route theories.

Models derived from the single-route theory were more successful, which is also supported by the findings presented in Tables 5–8. According to Table 5, in the answers of BM, total scores of the gender-based frequency outnumbered only slightly the mean value of cue validity, iconicity, and productivity ($p < .05$), whereas in the answers of MO, the difference was highly significant ($p < .001$). This means that both BM and MO followed the principle of the gender-based frequency significantly more often than not frequency-based principles, but MO had a better command of it than BM. This becomes even more obvious when the general frequency is added to cue validity, iconicity, and productivity and contrasted with the gender-based frequency: The values of the gender-based frequency in the MO group were still highly significant (Table 5), but the values of the BM group lost their significance. Highly significant results were obtained in both groups when the general frequency was contrasted with cue validity, productivity, and iconicity. Consequently, both BM and MO utilized plural allomorph distributions suggested by general and gender-based frequency, but BM used the gender-based frequency only

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

slightly more often than not frequency-based principles, probably because the gender-based frequency represents a cognitive challenge and demands a longer period of language contact. In contrast, MO actively used both kinds of frequency.

The plural marker characteristic most clearly associated with the total number of correct answers in the language tests *SETK 3-5* and *MSS b* was gender-based frequency. Children who have already mastered this characteristic scored higher in *SETK 3-5* and *MSS b* plural tasks as well as in *MSS b* subtests on grammar, vocabulary, word and sentence repetition (phonological short-term memory), articulation (low correlation in this case), speech comprehension, and *MSS b* total score in comparison with children who mastered general (not gender-based) frequency only. Also, command of the gender-based plural marker frequency was closely associated with a higher number of actively used plural allomorphs, a lower number of zero forms, and a good command of gender. There was also a clear negative correlation between the command of this plural allomorph characteristic and the classification of children as needing educational and/or medical help in acquiring/learning German, which means that children who had already acquired the gender-specific distribution of plural markers tended to be classified as not needing any help in acquiring/learning German. It remains open whether the gender-based frequency is indeed more closely associated with command of the German language than other plural allomorph characteristics, or whether it was better assessed due to its clear hierarchies of the most expected plural allomorphs. No such hierarchies exist for cue validity, productivity, and iconicity, or, rather, a plethora of definitions of these terms provides a broad spectrum of such hierarchies.

Plural formation in both groups – MO and BM – corresponded to a large extent to the principles of iconicity and cue validity, both predicting the dominance of the suffixes *-e*, *-s*, and *-(e)n* in the overgeneralizations. More complicated principles of the productivity were often ignored by the children in our sample. Although both MO and BM, that is, both linguistically more and less proficient groups, followed the same overgeneralization patterns, the linguistically less proficient children tended to use the plural markers which dominate in the input, *-(e)n* and *-e*, whereas the more proficient children used the comparatively rare plural suffix *-s* that has the highest ranking on the productivity scale. The dominance of *-s* is a sign of transition from the simple frequency-based phonetic–phonological patterns to the more complicated ones, which are, all in all, also based on frequency.

To a certain extent, all our overgeneralization models violate the principle of productivity in forms like *Fisches* (*-s* after sibilants), *Plartes* (*-s* after a schwa), *Bilden* (*-(e)n* with neuter nouns), and *Schiffer* (*-er* is not productive). Even the best models – 'Feminine vs. non-feminine nouns', 'Gender neglect', 'Gender neglect + schwa', and 'Gender neglect + -e' – contain overgeneralization patterns

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

contradicting various regularities referring to the productivity concept: *Plartes*, *Ribanes*, *Fisches*, *Buchen*. The regularities based on plural allomorph frequency were never violated in our models because only this plural allomorph characteristic provides unambiguous hierarchies of the most expected plural forms. The violation of the principle of cue validity is reflected only in the form *Schiffer*. Apart from the form *Gäbele* in the 'Gender only' model, all the proposed models of regularities behind overgeneralization patterns used highly iconic plural markers. *Gäbele* was also a violation of the schwa deletion rule, which is never violated in the target language. Hence, every model was to a certain extent a compromise between various plural allomorph characteristics and none of them can be considered perfect on all hierarchies.

It would be a mistake to propose highly complicated models to explain the overgeneralization schemata lying behind the error patterns of the children who have just begun to acquire/learn German. As they do not have enough input data, their analyses are simple, the regularities they extract from the input refer to the most obvious rules or tendencies and do not systematically take grammar categories into account because notions like gender have not yet been internalized. In this respect, the regularity '*-n* after *-e*' (*Rabe* 'raven' > *Raben*), where the last vowel *-e* does not have any morphological meaning, is of the same nature as the regularity '*-(e)n* after *-heit*' (*Freiheit* 'freedom' > *Freiheiten*), where the plural suffix is prescribed by morphology unambiguously. At the first stages of plural acquisition, the regularities behind the pluralization patterns can be boiled down to frequency plus phonetic–phonological patterns without any further motivation or differentiation. Later stages are characterized by a more and more versatile consideration of other regularities, which are partly based on grammatical and semantic aspects.

The transition from the simple frequency and phonetic–phonologically based patterns to the more complicated ones is also the transition to analogous formations. For instance, one important finding by MacWhinney (1978), replicated in our study, is the growing rate of analogous forms in comparison with zero and 'emergency' forms (default-like plural markers), abundant at the first stages of plural acquisition. Analogous plural forms of nonce words are, according to MacWhinney, those which correspond in their plural markers to the real words with similar phonetic–phonological structure: *Gände* like *Hände* 'hands'. Indeed, in the *SETK 3-5* data, the linguistically less proficient BM group preferred both zero forms (74% of all classifiable markers out of 1,399 in the answers of BM, 69% out of 1,883 in the answers of MO) and the plural marker *-(e)n* which gave way to other more versatile patterns dependent on more elaborate principles in the linguistically proficient group of MO (BM used their emergency plural, *-(e)n*, in 51% of all overgeneralizations, MO used their emergency plural, *-s*, in 41% of all overgeneralizations). An increasing number of

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

analogous forms can be linked to the growing vocabulary of children acquiring/learning German.

Although results of the study provide more evidence in favor of the single-route models than in favor of the dual-route ones, the concepts of cue validity, productivity, and iconicity of plural allomorphs can account for the overgeneralizations registered in our data only within certain limitations, with none of them definitely playing a major role. One of the most obvious limitations is the fact that the highly iconic plural marker *-er* was almost never encountered in the overgeneralizations, probably due to the high frequency of identical word-final sounds in words like *der Hammer* 'hammer', and therefore due to its low cue validity. Even to the correct *er*-forms, the more 'convincing' marker *-(e)n* was sometimes added (e.g., *Büchern* 'books'; 33 occurrences in *SETK 3-5*). The high validity of the plural marker *-e* plus umlaut is not reflected in the high frequency of the overgeneralizations (in fact, this plural allomorph was almost never overgeneralized), which could theoretically be traced back to its low scores on iconicity. Furthermore, both linguistically more and less proficient children, that is, MO and BM, tended to ignore the low productivity of neuter nouns with the plural marker *-(e)n* and to overgeneralize this very marker with the *SETK 3-5* items *Schiff* 'ship', *Buch* 'book', and *Bild* 'picture' more often than any other one. The plural allomorph *-e*, which is supposed to be productive with masculine and neuter nouns, dominated in our data also with any feminine noun which requires the plural allomorph *-e* plus umlaut.

Only plural markers with high rankings on at least three of the scales – frequency, cue validity, iconicity, and productivity of plural allomorphs – were widely used in the overgeneralizations. The plural allomorphs *-(e)n* and *-e* are high on the first three scales, *-s* is high on the last three. Thus, statements that highly iconic plural markers often substitute less iconic plural markers make much less sense than statements that plural markers with high rankings on CUE STRENGTH (comprising type and token frequency, cue validity, iconicity, and SALIENCE) do (Ewers, 1999). The frequency of nouns also plays a certain role, because more frequent items in *SETK 3-5* were on average produced correctly significantly more often than less frequent ones (Zaretsky et al., 2011).

Concepts like cue validity and productivity as factors in plural acquisition are so closely linked to frequency that they can be referred to as specific aspects of the frequency factor. Cue validity is nothing but the frequency of the plural allomorph in its function as the plural allomorph. The productivity is the frequency of the plural allomorph in combination with certain phonetic–phonological environments or, at later stages, with certain morphological or semantic features. Some conflicts between probabilistic patterns inevitably emerge, for instance, between the high frequency of *-e*, especially with nouns

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

ending in a consonant, and the low frequency of *-e* with feminine nouns. Yet for children who are still hardly aware of the concept of gender, it seems to be clear that *-e* must be overgeneralized with feminine items like *Hand* 'hand' because it ends in a consonant. The development of the pluralization patterns is thus the development from simple frequency-based patterns to the more and more complicated frequency-based ones.

The preference for *s*-plurals accentuated in the works of proponents of the dual-route models can be accounted for within the single-route framework as the preference for an iconic, productive plural marker with a high ranking on cue validity and closely associated with all three genders (or both groups of feminine and non-feminine nouns) in linguistically proficient children. There is no evidence for a universal compatibility of the plural marker *-s* with all phonetic–phonological environments, and thus for its default status. In this respect, our data do not support the dual-route model as it was described in the 'Introduction'.

A considerable subgroup of the sample, mostly BM, did not produce any correct or wrong plural forms at all in *MSS b*. Instead, they repeated singular forms. Because these children scored low in other language tasks as well, it is to be assumed that the critical mass of the language input was not yet accumulated by this subgroup, and no schemata or pluralization rules were internalized. The language deficits of these children might be linked to disabilities or impairments which influence language acquisition. This assumption is supported by the fact that, according to the classification of the language experts (speech and language pathologists, researchers), 82% of children who repeated all three singular forms in *MSS b* needed educational and/or medical support in learning German (cf. 31% for children who produced at least one correct or wrong plural form).

It can be assumed that BM use qualitatively different overgeneralization patterns. However, Zaretsky et al. (2013a) showed that MO and BM from our sample, irrespective of their national background, basically tended to apply similar pluralization hypotheses resulting in identical error patterns such as *-e* > *-(e)n* (e.g., *Fische* > *Fischen* 'fish'). Only the proportions of *s-*, *(e)n-*, and zero forms differed. The BM in our sample were therefore a rather homogeneous group, as far as the use of plural forms is concerned, and could be compared with the MO without subdivision into smaller groups. Hence, no such subdivision was utilized.

We are not aware of other studies with comparable study designs where several models of regularities behind overgeneralization patterns were tried out by means of the Wilcoxon or *t*-test for paired groups, these groups being total numbers of expected and/or unexpected answers. Hence, only certain aspects of our study are comparable with previous research. Among other things, the overgeneralization patterns reflected in Table 2 (*-e* > *-(e)n*, umlaut +

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

-e > *-e*, *-er* > *-(e)n*, *-(e)n* > *-s*, umlaut > *-(e)n*, *-s* > *-(e)n*) correspond to those found by Schaner-Wolles (2001), with only one exception: *-(e)n* > *-e* instead of *-(e)n* > *-s* in our data. Schaner-Wolles tested 76 monolingual Germans from Vienna aged 2;3 to 6;4 cross-sectionally with elicitation tasks which included eight items from our item set. Indeed, very little evidence can be found in the previous research for the overgeneralization pattern *-(e)n* > *-s*, probably due to its low frequency: *-(e)n* occurs mostly in nouns ending in a schwa and this combination (schwa + *-(e)n*) is one of the simplest pluralization rules in German (Mugdan, 1977). Consequently, the percentage of wrong answers for such items is comparatively low (cf. Zaretsky et al., 2013a). However, in a large plural elicitation study by Indefrey (2002), where only nouns ending in a schwa were utilized, this pattern dominated in the children's error patterns, together with zero forms, both in real and nonce words.

One of the most efficient models both for the MO and BM group was 'Gender neglect'. The model stressed the simplest phonetic–phonological regularities of the German language and neglected the distribution of plural markers governed by the category of grammatical gender. The model was partly based on the regularities of the plural formation described by Mugdan (1977), see Section 2 'Plural formation in German'. Table 2 confirmed these tendencies for our sample and thus also eight overgeneralization principles (or rules) chosen for the model 'Gender neglect'. According to Table 2, all items, except those with full vowels or schwa as final sounds, were incorrectly built either by repeating the singular or by overgeneralizing one or more plural allomorphs. Items following the two easiest rules, however, were merely repeated in more than 90% of the cases and were rarely substituted by other plural markers, presumably because children did not find evidence in their language input for other plural formation patterns and closely associated certain plural allomorphs with the word-final sounds (schwa and full vowels); see our rules (4) and (5) in the 'Gender neglect' model. Furthermore, items ending in a schwa followed by a consonant were clearly associated with *-n* and *-s*, that is, two non-syllabic suffixes, as was predicted by Mugdan (1977). Thus, three out of four of his principles found clear reflection in our data. Furthermore, items demanding *-e* plus umlaut are pluralized in Table 2 by adding *-e* alone (cf. our rule 6), and items demanding *-(e)n* are pluralized by *-s* (cf. our rule 2). With all *SETK 3-5* items, zero forms were the most widespread answer (cf. our rule 8). These tendencies were also mentioned by Mugdan (1977) as the most obvious ones in the German plural system.

Overgeneralization patterns and hence the frequency of certain plural allomorphs in the answers of children depend to a certain extent on the choice of the test items. In Table 3, not statistically significant results for the MO with respect to the models 'Feminine vs. non-feminine nouns' and 'Gender neglect' without considering nonce words became significant when

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

the nonce words were included in the calculations. This probably demonstrates that children reproduced memorized plural forms of real words and that some tendencies can be revealed only by the consideration of nonce items as well. Furthermore, our results revealed that items ending in a consonant increase the number of *e*-overgeneralizations, and items ending in a syllable with a schwa decrease this number and promote *n*- and *s*-overgeneralizations.

Although the overgeneralization patterns revealed in different studies could partly depend on the choice of items (nonce words vs. real words, seldom vs. highly frequent ones), a certain common denominator could be expected, which probably would demonstrate the tendency of a transition from default-like emergency plurals at the early stages to a more diverse system in the more proficient groups. Indeed, a number of spontaneous speech studies (Behrens, 2001; Bittner, 2000; Bittner & Köpcke 2001; Köpcke, 1998; Park, 1977; Szagun, 2001) demonstrated that the overgeneralizations of *-s*, *-(e)n*, and *-e* are ubiquitous and *-(e)n* dominates at the early stages, which is attributed to its frequency and productivity by some authors (e.g., Kauschke et al., 2011). The ubiquity of *-(e)n* among early plural forms led some researchers to believe that only *-(e)n* is used regularly at the preschool age, whereas the rules regulating the use of other plural markers are acquired later (Mills, 1985). In elicitation studies (Ewers, 1999; Laaha et al., 2006; MacWhinney, 1978; Mudgan, 1977; Niedeggen-Bartke, 1999; Schaner-Wolles, 1989; Spreng, 2004), the same three plural markers dominated, yet without the clear preponderance of *(e)n*-overgeneralizations. The comment by Kauschke et al. (2011), that the reason for the dominance of *-(e)n* in spontaneous speech and the high frequency of *-s* and *-e* in the elicitation tasks could be traced back to the nature of the tasks, must be revised, because in almost all cases, the children in the spontaneous speech studies compared were below the age of four, whereas in almost all elicitation studies they were aged four and above. Thus, the findings of both types of studies correspond to our data.

It can be concluded that the non-consideration of gender aspects in favor of phonetic–phonological rules of plural formation explains the patterns found in the answers of BM, but not those of MO, who rather consider the distinction between feminine and non-feminine nouns in combination with easy phonetic–phonological rules. Therefore, there is no common model which would perfectly explain the error patterns of both MO and BM. Although the models 'Gender neglect' and 'Gender neglect + -e' proved to be highly effective for the whole sample ($ps < .001$), the 'Gender neglect' model was only moderately successful in predicting the error patterns of BM and the 'Gender neglect + -e' model in predicting the error patterns of MO ($ps < .01$). This does not mean that the error patterns of BM and MO were qualitatively different: as was shown in Zaretsky et al. (2013b), the error patterns of older

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

BM corresponded to the error patterns of younger MO and, consequently, reflect earlier stages of language acquisition. Rather, it means that the error patterns of both BM and MO children are governed by predominantly frequency-based principles described in the single-route theory, and that no default plural marker providing evidence for the dual-route theory was identified.

REFERENCES

Bartke, S. (1998). *Experimentelle Studien zur Flexion und Wortbildung* [Experimental studies on inflexion and word formation]. Tübingen: Niemeyer.

Behrens, H. (2001). How to learn a minority default: the acquisition of the German -s plural. Unpublished manuscript. Leipzig: Max-Planck-Institute for Evolutionary Anthropology.

Bittner, D. (2000). Sprachwandel durch Spracherwerb? Pluralerwerb [Language change through language acquisition? Plural acquisition]. In A. Bittner, , D. Bittner, & K.-M. Köpcke (Eds.), *Angemessene Strukturen: Systemorganisation in Phonologie, Morphologie und Syntax* (pp. 123–140). Hildesheim: Olms.

Bittner, D., & Köpcke, K.-M. (2001). The acquisition of German plural markings: a case study in natural and cognitive morphology. In C. Schaner-Wolles, J. Rennison, & F. Neubarth (Eds.), *Naturally! Linguistic studies in honour of Wolfgang Ulrich Dressler presented on the occasion of his 60th birthday* (pp. 47–58). Turin: Rosenberg & Sellier.

Bybee, J. (1985). *Morphology: a study of the relation between meaning and form*. Amsterdam: Benjamins.

Bybee, J. (1988). Morphology as lexical organization. In M. Hammond & M. Noonan (Eds.), *Theoretical morphology* (pp. 119–141). San Diego, CA: Academic Press.

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**, 225–255.

Dressler, W. U., Mayerthaler, W., Panagl, O., & Wurzel, W. U. (1987). *Leitmotifs in natural morphology*. Amsterdam: John Benjamins Publishing Company.

Euler, H. A., Holler-Zittlau, I., van Minnen, S., Sick, U., Dux, W., Zaretsky, Y., & Neumann, K. (2010). Psychometrische Gütekriterien eines Kurztests zur Erfassung des Sprachstandes vierjähriger Kinder [Psychometric quality criteria of a language screening for four-year-old children]. *HNO*, **58**, 1116–1123.

Ewers, H. (1999). Schemata im mentalen Lexikon: Empirische Untersuchungen zum Erwerb der deutschen Pluralbildung [Schemata in the mental lexicon: empirical studies on the acquisition of the German plural system]. In J. Meibauer & M. Rothweiler (Eds.), *Das Lexikon im Spracherwerb* (pp. 106–127). Tübingen: Francke.

Gawlitzek-Maiwald, I. (1994). How do children cope with variation in the input? The case of German plurals and compounding. In R. Tracy & E. Lattey (Eds.), *How tolerant is universal grammar? Essays on language learnability and language variation* (pp. 225–266). *Tübingen*: Max Niemeyer Verlag.

Grimm, H. (2001). *SETK 3-5: Sprachentwicklungstest für drei- bis fünfjährige Kinder* [SETK 3-5: Test of language development for three- to five-year-old children]. Göttingen: Hogrefe.

Indefrey, P. (2002). Listen und Regeln: *Der Erwerb und Repräsentation der schwachen Substantivdeklination des Deutschen* [Lists and rules: acquisition and representation of the weak declination of German nouns]. Unpublished doctoral dissertation. Heinrich-Heine-University, Düsseldorf, Germany.

Institut für Deutsche Sprache (2009). *Korpusbasierte Grundformenliste DeReWo, v-40000g-2009-12-31-0.1, mit Benutzerdokumentation* [Corpus-based lexeme list DeReWo, v-40000g-2009-12-31-0.1, with a manual]. Mannheim: Institut für Deutsche Sprache, Programmbereich Korpuslinguistik. Online: <http://www.ids-mannheim.de/kl/projekte/methoden/derewo.html>.

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

Kauschke, C., Kurth, A., & Domahs, U. (2011). Acquisition of German noun plurals in typically developing children and children with specific language impairment. *Child Development Research*, DOI: 10.1155/2011/718925. Online: <http://www.hindawi.com/journals/cdr/2011/718925/>.

Köpcke, K.-M. (1993). *Schemata bei der Pluralbildung im Deutschen* [Schemata in the German plural formation]. Tübingen: Narr.

Köpcke, K.-M. (1998). The acquisition of plural marking in English and German revisited: schemata versus rules. *Journal of Child Language*, **25**, 293–319.

Korecky-Kröll, K., & Dressler, U. W. (2009). The acquisition of number and case in Austrian German nouns. In U. Stephany & M. D. Voeikova (Eds.), *Development of nominal inflection in first language acquisition: a cross-linguistic perspective* (pp. 265–302). Berlin: Walter de Gruyter.

Laaha, S., Ravid, D., Korecky-Kröll, K., Laaha, G., & Dressler, W. U. (2006). Early noun plurals in German: regularity, productivity or default? *Journal of Child Language*, **33**, 271–302.

Leipzig University, Department for Natural Language Processing (2010). *Wortschatz* [Vocabulary]. Online: <http://corpora.uni-leipzig.de/download.html>.

MacWhinney, B. (1978). The acquisition of morphophonology. *Monographs of the Society for Research in Child Development*, **174**.

Marcus, G., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: the exception that proves the rule. *Cognitive Psychology*, **29**, 189–256.

Mayerthaler, W. (1981). *Morphologische Natürlichkeit* [Morphological naturalness]. Wiesbaden: Athenaion.

Mills, A. E. (1985). The acquisition of German. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition. Vol. 1: The data* (pp. 141–254). Hillsdale, MI: Lawrence Erlbaum.

Mugdan, J. (1977). *Flexionsmorphologie und Psycholinguistik* [Flexion morphology and psycholinguistics]. Tübingen: Günther Narr.

Neumann, K., Holler-Zittlau, I., van Minnen, S., Sick, U., Zaretsky, Y., & Euler, H. A. (2011). *Katzengoldstandards in der Sprachstandserfassung: Sensitivität und Spezifität des Kindersprachscreenings (KiSS)* [Fool's gold standards in language screening: sensitivity and specificity of the Hessian child language screening test Kindersprachscreening (KiSS)]. *HNO*, **59**, 97–109.

Niedeggen-Bartke, S. (1999). Flexion und Wortbildung im Spracherwerb [Inflexion and word formation in the language acquisition]. In J. Meibauer & M. Rothweiler (Eds.), *Das Lexikon im Spracherwerb* (pp. 208–228). Tübingen: Francke.

Park, T.-Z. (1977). Plurals in child speech. *Journal of Child Language*, **5**, 237–250.

Phillips, B. S., & Lowell, B. (1980). The acquisition of German plurals in native children and nonnative adults. *International Review of Applied Linguistics in Language Teaching*, **18**, 21–29.

Pinker, S. (1999). *Words and rules: the ingredients of language*. London: Weidenfeld & Nicolson.

Schaner-Wolles, C. (1989). Plural- vs. Komparativerwerb im Deutschen – Von der Diskrepanz zwischen konzeptueller und morphologischer Entwicklung [Plural vs. comparative acquisition: on the discrepancy between conceptual and morphological development]. In H. Günther (Ed.), *Experimentelle Studien zur deutschen Flexionsmorphologie* (pp. 155–186). Hamburg: Buske.

Schaner-Wolles, C. (2001). On the acquisition of noun plurals in German. In C. Schaner-Wolles, J. Rennison, & F. Neubarth (Eds.), *Naturally! Linguistic studies in honour of Wolfgang Ulrich Dressler presented on the occasion of his 60th birthday* (pp. 451–460). Turin: Rosenberg & Sellier.

Spreng, B. (2004). Error patterns in the acquisition of German plural morphology: evidence for the relevance of grammatical gender as a cue. *Toronto Working Papers in Linguistics*, **23**, 147–172.

Stemberger, J. P. (1999). Frequency determines defaults in German: default perfect -t vs. irregular plural -s. *Behavioral and Brain Sciences*, **22**, 1040–1041.

Szagun, G. (2001). Learning different regularities: the acquisition of noun plurals by German-speaking children. *First Language*, **21**, 109–141.

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press

Vollmann, R., Sedlak, M., Müller, B., & Vassilakou, M. (1997). Early verb inflection and noun plural formation in four Austrian children: the demarcation of phases and interindividual variation. *Papers and Studies in Contrastive Linguistics*, **33**, 59–78.

Walter, S. (1975). *Zur Entwicklung morphologischer Strukturen bei Kindern* [On the development of morphological structures in children]. Unpublished master's thesis, Heidelberg University, Institute for Psychology.

Wegener, H. (1995). *Die Nominalflexion des Deutschen – verstanden als Lerngegenstand* [The nominal inflection in German as a study subject]. Tübingen: Max Niemeyer.

Wurzel, W. U. (1984). *Flexionsmorphologie und Natürlichkeit* [Inflexion morphology and naturalness]. Berlin: Akademie-Verlag.

Zaretsky, E., Euler, H. A., Neumann, K., & Lange, B. P. (2014). Sociolinguistic predictors of language deficits in pre-school children with and without immigrant background. In *The Asian Conference on Language Learning. Osaka, Japan 2014. Conference Proceedings* (pp. 42–53). Nagoya: International Academic Forum.

Zaretsky, Y., Lange, B. P., Euler, H. A., & Neumann, K. (2011). Pizzas, Pizzen, Pizze: frequency, iconicity, cue validity, and productivity in the plural acquisition of German preschoolers. *Acta Linguistica*, **2**, 22–35.

Zaretsky, E., Lange, B. P., Euler, H. A., & Neumann, K. (2013b). Acquisition of German pluralization rules in monolingual and multilingual children. *Studies in Second Language Learning and Teaching*, **3**, 551–580.

Zaretsky, E., Neumann, K., Euler, H. A., & Lange, B. P. (2013a). Pluralerwerb im Deutschen bei russisch- und türkischsprachigen Kindern im Vergleich mit anderen Migranten und monolingualen Muttersprachlern [Plural acquisition in German in Russian and Turkish speaking children in comparison with other immigrants and monolingual native speakers]. *Zeitschrift für Slawistik*, **58**, 43–71.

https://doi.org/10.1017/langcog.2014.51 Published online by Cambridge University Press