From formalism to intuition: probing the role of the trochee in German nominal plural forms in L1 and L2 German speakers

Katharina S. Schuhmann[1]* and Laura Catharine Smith[2]

[1]Department of German, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany,
[2]Department of German and Russian, Brigham Young University, Provo, UT, United States

Accounting for plural formation in Standard German (SG) nominals has proven to be a challenging endeavor. Numerous formalisms and models have been proposed and intensely debated over the past decades. The fundamental difficulty lies in the fact that German has a large number of suffix allomorphs, some of which can be used with or without stem-vowel fronting/raising (umlaut). Current research suggests that, at the segmental level, it is impossible to fully predict how plurality will be marked for a given singular form. At the suprasegmental level, however, the vast majority of German plurals, except plurals ending in <-s> /-s/, exhibit a specific prosodic shape word-finally: a strong-weak pattern, i.e., a sequence of a stressed syllable followed by an unstressed syllable. In other words, German plurals tend to end in a disyllabic trochee. Previous experimental investigations have sought to provide empirical evidence in favor of various formal models. To date, these experimental studies have focused primarily on the segmental composition of plural suffixes. It remains untested—and thus largely unknown—whether the prosodic pattern at the interface between morphology and phonology is an active, productive part of the grammar of first language (L1) and second language (L2) users of German across proficiency levels. We therefore set out to test whether users actively apply the trochaic principle in the production and comprehension of German plural nouns. To this end, we tested L1 German and L1 English-L2 German users across four proficiency levels on a non-word plural elicitation task, in which they produced plural forms for non-words, akin to a *wug*-test; L2 users additionally completed a plural elicitation task with existing German nouns. All users then participated in a grammatical acceptability judgment task, in which they rated German nouns with various incorrect and correct (i.e., SG) plural forms on a Likert scale. L2 learners produced more trochaic plural forms as proficiency increased, and more advanced users showed a stronger correlation between their ratings and plural forms depending on the forms' correctness and prosody. We further analyzed how prosodic patterns varied with morphological context across proficiency levels, before discussing how the data can be accounted for within various models of German plurals.

KEYWORDS

morphophonology, prosody, trochee, German plurals, L1, L2 development, allomorphs, schema

1 Introduction

Nominal plural inflection in Standard German (SG), henceforth German, is highly complex and has spurred many theoretical accounts over the past several decades (e.g., Augst, 1979; Bittner, 1991; Köpcke, 1988, 1993; Mugdan, 1977; Neef, 1998; Trommer, 2021; Wiese, 1996, 2009; Wunderlich, 1999; Wurzel, 1998; *among others*). The (ensuing) debates have attempted to explain the morphology and morphophonology of the exponents, i.e., markers, of German nominal plurals, and how these markers might be acquired. The fundamental difficulty lies in the fact that SG exhibits a large number of plural allomorphs, including six different suffix allomorphs, some of which can also be combined with stem vowel fronting and/or raising, i.e., umlaut,[1] U. As schematized in Table 1, German espouses nine phonetically distinct options for marking nominal plural formation when the suffix choices and the possibility of umlaut are combined, i.e., <-en>, <-n>, <-e>, <-er>, <-s>, -Ø, U+<-e>, U+<-er>, and U+-Ø, leaving aside exceptional cases, such as suppletion.

The pattern of nominal plural data in German in Table 1 presents a challenge for theoretical models as well as for questions of learnability and models of language acquisition. Indeed, it is impossible to fully predict how plurality will be marked for a given singular form at the segmental level. Nonetheless, linguists have identified co-occurrence restrictions between certain suffixes and umlaut, e.g., the suffixes <-n> /-n/, <-en> /-ən/, and <-s> /-s/ can never co-occur with umlaut (Augst, 1979; Wurzel, 1998; see also Trommer, 2021 for discussion, and Archibald, 2022 for an experimental L2 study). In addition, numerous tendencies have been noted regarding the type of plural marking a given singular form is likely to take, depending on features such as noun gender (e.g., feminine nouns are more likely to end in *-(e)n* or the word-final phonological composition of the singular form (e.g., no explicit ending for masculine and neuter nouns ending in *-el, -en,* or *-er*) (see Section 2 below). At the suprasegmental level, however, the vast majority of German plurals, except those ending in <-s> /-s/, share a specific prosodic shape, namely, they overwhelmingly end in a strong-weak syllable sequence ($\sigma_{strong}\sigma_{weak}$), i.e., a sequence of a stressed and an unstressed, reduced syllable, often analyzed as a disyllabic trochaic foot (e.g., Neef, 1998; Smith, 2004, 2020, 2022; Wiese, 1996, 2009).

Descriptive accounts of German nominal plurals that take syllabic or prosodic structure into account typically agree on this predominant strong-weak disyllabic structure word-finally. Theoretical accounts, however, differ in whether and how they capture this metrical pattern in their formal models (see Section 2). Indeed, most theories and experimental investigations have focused on the segmental composition of plural markers, i.e., the distribution and patterning of the plural suffix exponents, often leaving aside the role of prosody. In this study, we focus on the prosodic structure and whether it might play an active role in the grammar of first language (L1) and second language (L2) German users with regard to plural formation. To this end, we tested L1

and L2 German speakers to examine whether their production and perception of plural forms provide experimental support for the role of prosody in nominal plural formation. For the L2 German users, we present data from four different proficiency levels in a cross-sectional design to test whether and how the production and perception of the prosodic pattern in plural forms changes with increasing language experience. This design allows us to test whether L2 users show the target-like prosodic pattern from the start, and to trace their acquisitional path. A longitudinal or cross-sectional study enables us to examine whether this prosodic pattern emerges suddenly, all-at-once, or whether it develops progressively as learners' language competencies grow.

This article thus presents behavioral, psycholinguistic data that are in line with formal models of German plurals in which the grammar of L1 German speakers, and with increasing proficiency adult L1 English-L2 German learners, contains a prosodic condition for German nominal plural formation. Although users are not explicitly aware of this prosodic requirement, the results of this study demonstrate an overwhelming tendency for plurals to end in a disyllabic trochee, i.e., a sequence of stressed-unstressed syllables word-finally. This prosodic pattern is shown for speakers' productions and in their judgments of the well-formedness of correct (trochaic) and incorrect (trochaic and non-trochaic) plural forms where L1 and—to varying degrees based on proficiency—L2 users rate trochaic forms as more well-formed compared to non-trochaic forms.

This paper is organized as follows. Section 2 provides a brief review of previous theoretical research on German plural formation. Next, Section 3 presents the methodology of the study which tests L1 and L2 German users' prosodic intuitions for German plural formation in (a) a non-word plural elicitation task (and, for L2 speakers only, also a plural elicitation task with existing German nouns), and (b) a well-formedness judgment task. In Section 4, we present the results of the study, including how prosody and morphology interact. Based on these findings, Section 5 answers the research questions and addresses theoretical implications, limitations, and open questions. The paper concludes in Section 6.

2 Background and literature review

2.1 German plurals: rule-based approaches and tendencies

As noted, accounting for German nominal plural inflection is challenging not only because of the large number of plural allomorphs—several suffix allomorphs used with or without stem vowel fronting and/or raising (umlaut)—but also because it does not seem possible to fully predict how a given singular will be marked for plural. Nevertheless, a number of studies have revealed tendencies for nouns to select their plural marker based on specific features and characteristics of words. These features include the gender of a noun and word-final phonological characteristics in the singular form (Augst, 1979; Mugdan, 1977; Köpcke, 1988; Wurzel, 1984; Duden-Grammatik, 1995), especially with respect to the level of sonority of the segments (e.g., Laaha, 2011). Köpcke (1988) also discusses semantic features such as animacy. However, these

---

1 Note that only the vowels represented by <a, o, u> and the diphthong <au> are able to form an umlaut: /aː, a, oː, ɔ, uː, ʊ, aʊ/ become /ɛː, ɛ, øː, œ, yː, ʏ, ɔɪ/, respectively.

TABLE 1 Overview of nominal plural allomorphs in SG.

Affix type	Umlaut pattern	Example without umlaut	Example with umlaut
-e [ə]	variable (w/ or w/o umlaut)	der Hund—Hund-**e** (*the dog, dogs*)	der Hut—H**ü**t-**e** [u]——[y] (*the hat, hats*)
-er [ɐ]	always umlaut (U-able V)	das Kind—Kind-**er** (*the child, children*)	das Blatt—Bl**ä**tt-**er** [a]——[ɛ] (*the piece of paper, papers*)
-en [ən, n̩]	never umlaut	die Frau—Frau-**en** (*the woman, women*)	
-n [n]	never umlaut	die Tasse—Tasse-**n** (*the cup, cups*)	
-Ø (*no overt ending*)	variable (w/ or w/o umlaut)	der Keller—Keller-Ø (*the basement, basements*)	der Apfel—**Ä**pfel-Ø [a]——[ɛ] (*the apple, apples*)
-s [s]	never umlaut	das Kino—Kinos (*the movie theater, movie theaters*)	

Grayed out cells indicate non-existent plural formations.

tendencies do not always denote the pluralization of a noun in a "deterministic" way that applies in all cases. Most of the time, these tendencies are more "probabilistic" in nature (Szagun, 2001). Examples of deterministic plural rules include (based on Szagun, 2001, *inter alia*):

a) Nouns ending in <-e> ([ə]) form the plural with the suffix <-n> /-n/, e.g., *der/die Bote-n* "messenger-s", *die/die Blume-n* "flower-s", *das/die Auge-n* "eye-s"

b) Non-feminine nouns ending in *-el, -en, -er, -chen*, and *-lein* take the null-suffix -Ø, e.g., *der/die Eimer-Ø* "bucket-s", *das/die Steuer-Ø* "steering wheel-s" (cf. feminine *die/die Steuer-n* "taxes")

c) Nouns ending in a syllable with an unstressed full vowel select for the plural suffix <-s> /-s/, e.g., *der/die Park-s* "park-s", *die/die Oma-s* "grandma-s", *das/die Auto-s* "car-s".

More commonly, however, these so-called "rules" simply express tendencies that hold only with a certain probabilistic likelihood. For instance, feminine nouns form the plural with <-n> /-n/ in 73% of cases, regardless of the word-final segmental characteristics, e.g., *Leiter-n* "ladder-s",[2] while masculine and neuter nouns form the plural with <-n> /-n/ or <-en> /-ən/ in 9% and 4% of cases, respectively (Duden-Grammatik, 1995, see also Szagun, 2001). Thus, plural allomorphy patterns—including the distribution of gender-based plural exponents—show tendencies rather than rules.

2.2 German plurals and prosodic structure

One additional and crucial tendency, however, is often overlooked in general descriptions of German plurals despite a widespread agreement among researchers. It is the observation

that syllables and their organization into feet, i.e., metrical or prosodic structures, also play a role in German plural formation (e.g., Eisenberg, 2020; Köpcke, 1988, 1993; Laaha, 2011; Smith, 2020; Wegener, 1995; cf. Binanzer and Wecker, 2020). Indeed, the vast majority of German plurals, except plurals with the suffix <-s> /-s/ (see further details below), end in the specific prosodic shape illustrated in Table 1, namely a word-final sequence of a stressed and unstressed reduced syllable (e.g., Eisenberg, 2020; Neef, 1998; Smith, 2004, 2020; Wiese, 1996, 2009; *among many others*). Assuming that syllables are organized into feet, this generalization can be expressed as plural forms predominantly showing a word-final disyllabic trochee (see below for formal analyses). This generalization holds for plural nouns regardless of the shape of their singular form. For example, all nominal plural forms in Table 1 show a word-final trochee (the period "." marks syllable boundaries): *Hü.t-e* "hats", *Kin.d-er* "children", *Frau.-en* "women", and *Äp.fel-Ø* "apples". This is regardless of the underlying shape of the corresponding singular forms, e.g., non-trochaic singular *Hut* "hat" alongside trochaic *Ap.fel* "apple".

Although this is a strong prosodic *tendency*, there are exceptions. First, as noted above, plurals marked by <-s> /-s/ do not fall under this prosodic generalization; this means that *-s* plurals for singular forms that end in final trochee will show a final trochee in the plural, e.g., *Sto.ry-s* "stories", while those whose singular forms do not end in final trochee will not show a final trochee in the plural, e.g., *Klub-s* "clubs". For all other (monomorphemic) plurals formed using the suffix options in Table 1, the prosodic generalization captures most, albeit not all, plural forms. Consider one group of exceptional cases, namely singulars ending in a specific derivational suffix (e.g., *-ung, -chen, -keit/-heit*, etc.), which each require a specific plural suffix. For instance, words ending in *-ung* always take the plural suffix *-en*, as in *'Ta.gung.-en* "conference-s", thus yielding non-trochaic plural forms. Words ending in *-chen* always take the plural suffix -Ø, as in *'Mäd.chen-Ø* "girl-s" or *'Vö.gel.chen-Ø* "little bird-s", but may or may not result in a trochaic plural form. Moreover, the suffix *-in*, which marks animate objects as feminine, always requires the *-(n)en* plural ending, as in *'Kun.din.-nen* "fem. customer-s", a non-trochaic plural form.

---

2   It should be noted that nouns ending in <-e> ([ə]) always form the plural with <-n> /-n/, including feminine nouns ending in <-e> ([ə]).

While it may be argued that the disyllabic trochee is epiphenomenal for plural formation in German, a similar prosodic tendency, i.e., a preference for trochaic forms, has been found in other lexical and morphological classes in German, both historically and in modern times.[3] More generally, it has been proposed that German has a "preference for the trochaic foot" (Féry, 1994, p. 31), and that the trochee is the most important foot, followed—far behind—by the dactyl in German (Eisenberg, 2020). German, just as English, has been labeled a so-called "trochaic" language with regard to word stress, as native German words are most commonly stressed on the penultimate syllable. This can be explained by the fact that most native words have two syllables, of which the final one is reduced and unstressed (Domahs et al., 2014, p. 62).

Even beyond the tendency for German to be trochaic broadly speaking, Eisenberg (2020) proposes that the trochee is the most grammaticalized foot in both inflectional and derivational morphology. This tendency to draw on the trochee is a feature that German has inherited from its Germanic origins where entire lexical classes and morphological functions were reshaped in alignment with the trochaic foot, from high vowel deletion in Old English, Old High German, and Old Saxon nouns and verbs, to Old Frisian Vowel Balance (Smith, 2024). Likewise, the trochee has been found to shape plurals and diminutives in Dutch (Booij, 1998; van der Hulst and Kooij, 1998; Smith, 2004), as well as in various German dialects (Wiese, 2009; Smith, 2020). Similar findings are noted for other morphological processes including noun shortening, e.g., *Universität—Uni* "university" (e.g., Itô and Mester, 1997; Schuhmann, 2015). In all of these cases, the critical importance of the trochee (whether syllabic or moraic, i.e., based on syllable weight) in shaping these morphophonological processes is demonstrated by a greater prevalence of the trochee in these specific forms than found more generally in the lexicon.[4] Whether this is epiphenomenal or teleological in nature (for instance via templates) is beyond the scope of the current study. Nevertheless, these phenomena, historical and modern, highlight the persistent association of the trochee with morphological functions since the earliest records of German.

Reflecting on the modern German language, Féry (1994) likewise argues that a disyllabic trochee is built into both inflectional and derivational morphology whenever possible. To illustrate this, adjectival inflections are largely trochaic such that monosyllabic adjectives like 'green' become trochaic forms such as *'grü.n-e / 'grü.n-er / 'grü.n-es* "green, fem. or pl. / green masc. / green, neutr." in attributive position and indefinite contexts. In

definite contexts in attributive position, adjectives end in -*e*, e.g., trochaic *'grü.n-e*; all adjectival plural forms in attributive position end in -*en*, e.g., trochaic *'grü.n-en*. Note, however, that predicative adjectives are uninflected; thus, typical monosyllables like *'grün* remain non-trochaic in this context. Despite this predominance of trochaic forms in adjectival inflection, variation also exists. Disyllabic adjectives ending in -*er* and -*en* such as *'hei.ter* "cheerful" or *'troc.ken* "dry" allow both trochaic and dactylic inflected forms, e.g., *'hei.tres* and *'hei.te.res* "cheerful, neutr.", respectively (Eisenberg, 2020, p. 149).[5]

Stems have typically been described as trochaic in German. Yet, stems can be analyzed as trochaic either with respect to syllables (disyllabic) or moras (bimoraic), a different weight unit. Bimoraic stems are heavy, binary monosyllables (Féry, 1994). Thus, German stems can be said to be mostly trochaic if this allows for both disyllabic and bimoraic trochees. Another nominal inflection that shows some trochaic forms is genitive singular (Eisenberg, 2020). Yet, only weak masculine (and some neuter) nouns such as *Mensch* "human" require a trochaic form for the genitive singular and the plural form, i.e., *des/die 'Men.sch-en* (Eisenberg, 2020, p. 149). Besides weak masculine nouns, only nouns in sibilants require trochaic genitive singular forms due to phonotactic reasons, e.g., *des Ti.sch-es* "table, gen. sgl." (Eisenberg, 2020, p. 149). Otherwise, genitive singular forms do not have to be trochaic—in contrast to the basic pluralization pattern.

In sum, German is an overwhelmingly trochaic language for which the trochee has been analyzed as an unmarked prosodic constituent (Féry, 1994). While many lexical forms within inflectional and derivational morphology—in addition to plain stems—can be described as trochaic, the disyllabic trochee has been considered a "necessary condition" ("notwendige Bedingung") only for plurals (Eisenberg, 2020, p. 149). Overall, this indicates that the prosodic trochaic pattern in the plural formation of German nouns is firmly grammaticalized and applies even more consistently than in other sub-parts of the German language system.

2.3 The importance of prosody in L1 and L2 acquisition

Work on the acquisition of German plural formation by children with specific language impairment (SLI) has shed light on the role of prosodic requirements for plural formation and

---

3 In German, the trochee can take the form of bimoraic monosyllables, e.g., *Heu* "hay", or disyllables, e.g., *'Va.ter* "father" (Féry, 1994, p. 1). Of these two possible trochees, only the disyllabic trochee is associated with German plurals.

4 Indeed, McCarthy and Prince's (1996) work on Prosodic Morphology first made the connection between prosodic units like the foot in a variety of morphological processes, demonstrating that the foot can be influential not only in stress placement, but in shaping morphological patterns in a language. Psycholinguistic studies such as those discussed in Section 2.3 substantiate this intersection between prosody and morphological processes outlined in various theoretical approaches.

---

5 Similar patterns can be seen in possessive determiners. For instance, when the inflectional suffix -*e* is added to *'eu.er* "your, pl. non-fem", the resulting inflected form is trochaic, i.e., *'eu.re* "your, fem. sg./pl." Just as with adjectives, variation can be seen with certain forms. For example, when *'eu.er* or *'un.ser* "our" receive the inflectional suffixes -*e*, -*es*, or -*er*, the resulting inflected forms may or may not include the stem-final -*e*-: both *'eu.e.re, 'eu.re* and *'un.se.re, 'uns.re* are possible. With the inflectional suffixes -*em* and -*en*, either the stem-final or the suffix -*e*- might not surface in the inflected forms, e.g., *'un.se.ren / 'uns.ren / 'un.sern* (Duden-Grammatik, 1998, p. 336). Thus, certain inflected possessive determiners also show variation between trochaic and dactylic forms. Eisenberg (2020) further proposes that a dactyl occurs in inflected comparative forms such as *'klei.ne.ren*, in which an -*en* inflectional suffix is added to the comparative form *'klei.ner* "small, comp."

plural acquisition (Kauschke et al., 2013) and the role of prosody in morphosyntactic difficulties more generally (Domahs et al., 2013). In a plural elicitation task with existing and non-words, Kauschke et al. (2013) showed that German-speaking children with SLI (mean age 7.5) made more mistakes when forming plurals and produced fewer "prosodically optimal" plurals compared to age- and vocabulary-matched children without SLI. These findings underscore that the children with SLI had reduced sensitivities for the prosodic requirements of German plural marking. Kauschke et al. (2013, p. 574) argue that morphological processes are influenced by prosody, and that even more fundamentally, "morphological acquisition may in part be linked to the fact that children's acquisition of grammatical morphemes is closely tied to the development of prosodic representations (Demuth, 2009)." In another clinical study with a German-speaking patient showing primary progressive aphasia, Domahs et al. (2017) report that the patient was more likely to interpret a stimulus as a plural form if it included any of the cues of Köpcke's schema model (discussed below) and if it conformed to the "optimal prosodic plural form" (Domahs et al., 2017, p. 206, *translation KS*). Such studies point to the critical role of prosody in German for shaping morphological processes, representations, and acquisition, including plural formation.

The overall importance of prosody in L1 acquisition of morphology has been shown cross-linguistically. Demuth (2009) synthesizes cross-linguistic research on child L1 acquisition of prosodic morphology in support of the so-called Prosodic Licensing Hypothesis. The findings support the notion that frequently reported variability in the production of grammatical morphemes can be explained by the children's developing prosodic representations. For instance, Demuth (1994) argued that children acquiring the Bantu language Sesotho initially produce specific morphological prefixes only when they are footed into a disyllabic foot. It was proposed that children would first acquire morphemes prosodified as part of a foot, such as internal clitics. Further data from Gerken (1996) and Demuth and McCullough (2009) on children learning English support this approach, where it was shown that the children were more likely to produce determiners when they were part of a prosodic foot, i.e., prosodically licensed. Interestingly, cross-linguistic comparisons of L1 acquisition from English, German, Spanish, and Italian (Lleó and Demuth, 1999) show that determiners are produced earlier by children learning Romance languages than children learning Germanic languages. In Spanish, determiners are cliticized to the nouns as proclitics, while in German, a determiner is prosodified as a separate foot when it occurs in its full, unreduced, and un-cliticized form (Lleó and Demuth, 1999; Wiese, 1996). Thus, Lleo and Demuth's analyses provide further evidence for the proposal that the appearance of grammatical inflectional morphemes depends on prosodic licensing during child L1 acquisition.

Goad and White (2019) have put forth a similar proposal in which prosody constrains the acquisition of grammatical morphemes during L2 development. According to their Prosodic Transfer Hypothesis (PTH), L2 learners' non-target-like forms can be accounted for by the transfer of their L1 prosodic constraints. In their current revised PTH, they further argue that not only the production but also the comprehension and processing of L2 grammatical morphology is constrained by the transfer of learners' L1 prosodic structures to their L2.

2.4 Formal accounts and acquisition/ processing models of German plurals

A wide variety of theoretical analyses and models have been proposed to account for German nominal plural formation. A large number of accounts have focused on the plural allomorphs and restrictions on their co-occurrence with stem vowel changes, i.e., umlaut, while the role of prosody varies from "not discussed" to central. The following discussion presents an overview of the approaches, ranging from the debate about dual- vs. single-route models, to schematic approaches, and generative phonological analyses.

2.4.1 Models of L1 language acquisition and processing

In terms of modeling L1 acquisition, much of the debate about L1 German plural acquisition over the past two and a half decades has centered on whether the acquisition, processing, and use of the many German plural forms are best explained by a dual- or single-route model. The generativist perspective tends to favor an intrinsic, dual-mechanism perspective (Clahsen, 1999; Marcus et al., 1995), which is modeled on the inflectional system in English. This perspective suggests that, on the one hand, regular forms are generated and comprehended via a rule-based mechanism that combines or decomposes the word-stem and an affix ("stem+affix"). Irregular forms, on the other hand, are memorized and stored holistically as discrete lexical entries in the mental lexicon, but they "pattern according to phonological similarities" (Köpcke et al., 2021, p. 4). However, attempting to define what counts as the "regular" plural form in German, or whether there even is a "regular" plural form in German, is contentious. Meanwhile, analyses of English straightforwardly map the most frequent forms onto the "regular" forms.

For German, proponents of the dual-route model initially argued that the -*s* plural is the "regular" plural, which is, despite its low frequency of about 4% (Szagun, 2001) to 6% (Zaretsky et al., 2013), processed fastest among all plural markers in German. Dual-route proponents have also argued that the -*s* plural is the "regular" plural based on Clahsen's original claim that -*s* was the most frequently overgeneralized plural marker in children; this claim, however, has since been refuted (see Laaha et al., 2006, p. 277 for a brief overview). According to this dual-route model, all other plural forms are irregular plurals (Clahsen, 1999; Clahsen et al., 1992; Marcus et al., 1995; for empirical evidence, see Beretta et al., 2003; Clahsen et al., 1997; Sonnenstuhl et al., 1999; for an overview, see Clahsen et al., 2003; see also Köpcke et al., 2021). Slightly later versions of this dual-mechanism model concede that the -*n* plural for feminine nouns ending in schwa has a special status among all irregular plurals due to its highly predictable nature (Penke and Krause, 2002; Sonnenstuhl-Henning, 2010; cf. Köpcke et al., 2021, p. 4). Other approaches include a race-model variant of the dual-route view (e.g., Baayen and Schreuder, 1999; Schreuder and

Baayen, 1995; see also Pinker and Ullman, 2002; Laaha et al., 2006), which assumes that both a parsing and a retrieval route operate at the same time. Thus, according to such race-model variants, both memory and rule-application processes apply in parallel. The winning route depends on the characteristics of the individual inflected words, such as their frequency, their morphology, the transparency of their phonology and semantics, and their lexical neighborhood effects (cf. Laaha et al., 2006, p. 273).

Conversely, proponents of the single-route model argue that children construct grammatical structures from their input and that all forms are processed in the same way, via one mechanism rather than two. These include the schema-based models of Köpcke et al. (discussed below), constructivist theories, and connectionist models (see Köpcke et al., 2021, p. 5 for a brief overview). For instance, Hahn and Nakisa (2000) provide simulation and experimental data in support of single-route connectionist systems. Laaha et al. (2006) present results from a plural elicitation study with children that can be explained by either a single-route model or a race variant of the dual-route model: Rules that are more productive and transparent are more likely to be used than the lexical route (Laaha et al., 2006, p. 299). Their data also support their hypothesized role of the degree of productivity in the speed of acquisition of the plural form of a given noun.

This literature has typically focused on factors that constrain the acquisition of the various plural suffix options as outlined in Table 1 (i.e., *-e, -er, -en, -n, -Ø, -s*), with or without umlaut. To the best of our knowledge, so far, the prosodic pattern in German plurals has not played a central role in this debate of single- vs. dual-route models and their variants. However, syllabic or prosodic aspects are at least mentioned—or take center-stage—in the specific schematic and generative approaches discussed below.

2.4.2 Schematic approaches

One of the most prevalent models for German nominal plural formation is the "prototype-based 'schema' model" by Köpcke and colleagues (e.g., Köpcke, 1988, 1993, 1998; Köpcke et al., 2021; Köpcke and Wecker, 2017). In this schema-based model, specific plural forms signify plurality more robustly than others because certain features or cues, e.g., polysyllabicity, umlaut, the determiner *die* ("pl." or "fem. sg."), and word-final characteristics (see Köpcke, 1988 for particulars), or combinations of these cues, make a word appear more like a typical plural. Plural markers are analyzed as "abstract schemata" consisting of these cues as opposed to "individual morphemes" (Köpcke, 1988, p. 330). Arguing that language users engage in a "schema-matching process", Köpcke proposes that individual plural forms differ in how strongly they are associated with the function of plurality due to differences in the "cue strength" of each feature's salience, type frequency, cue validity, and iconicity. For instance, in the case of the plural markers introduced above, many of these markers—e.g., the suffixes <-e> /-ə/ and <-er> /-ɐ/—can also occur as word-final segments on singular forms, e.g., *Leh.**rer*** "teacher" or *Lö.**we*** "lion". It is argued that the link between form and function becomes stronger when a specific form occurs particularly often in a plural form (type frequency) while occurring rarely in a singular form (validity). To provide a concrete example, in this proposal, nouns ending in <-en> indicate plurality more strongly than nouns ending in <-e>.

This is because word-final <-en> occurs more frequently in the plural than in the singular, while singular forms ending in <-en> are uncommon. This model has recently been used to describe the acquisition of plural forms in L1- and L2-speaking children in German primary schools (Binanzer and Wecker, 2020; Köpcke and Wecker, 2017; Wecker, 2016). The latest instantiations of schematic approaches now also reflect the increased attention paid to the prosodic regularity of German plurals as they incorporate the trochee as the optimal prosodic form for plurals, rather than just polysyllabicity vs. monosyllabicity or number of syllables. This is the case in a clinical case-study with an individual with primary progressive aphasia (Domahs et al., 2017) and in a study on child L2 German acquisition among elementary school students in Germany (Köpcke and Wecker, 2017). This last study led to the most recent elaboration of the schema model (Köpcke and Wecker, 2017; Wecker, 2016) with "schema-pairs" (Köpcke et al., 2021, p. 7), which are intended to capture the paradigmatic relation between the two first-order schemas for singular and plural, respectively. As an example, a second-order "super"-schema (Köpcke and Wecker, 2017, p. 85) accounts for the connection between singular schemas with the feminine article and an <-e> ending, on the one hand, and plural schemas with the plural article and the <-en> suffix, on the other hand (Köpcke and Wecker, 2017, p. 85).

2.4.3 Generative analyses

To account specifically for the prosodic pattern evident in German plural formation, several formal accounts have been proposed from a generative linguistic perspective. In particular, early morphophonological formalisms describe the prosodic shaping of plural formation in German in terms of edge effects; regardless of the singular form, the right edge of the plural must align with a disyllabic trochee. Drawing on Prosodic Morphology (McCarthy and Prince, 1996, et seq.), Smith (2004, 2020) analyzes plurals using a prosodic, foot-based template that requires plurals to end in a disyllabic trochee, i.e., (σσ)#. In Smith's analysis, the right edge of plurals is mapped to a complex prosodic template consisting of a disyllabic trochee with a final schwa-syllable (a so-called "schwallable", Booij, 2002). Thus, if a singular form is already trochaic, e.g., *'Leh.rer* "teacher" or *'Lam.pe* "lamp", then the plural ending must be non-syllabic, i.e., -Ø or -*n*, respectively, so as to not disturb the trochee, thus *'Leh.rer-Ø* "teacher-s" or *'Lam.pe-n* "lamp-s". However, if the singular form does not end in a trochee, e.g., *Tag* "day", *I.'dee* "idea", or *Feld* "field", then the plural ending is syllabic, i.e., -*e, -en,* or -*er*, respectively (*'Ta.g-e* "day-s", *I.'de.-en* "idea-s", or *'Fel.d-er* "field-s"), to satisfy the trochaic template. While Smith views templates as tendencies rather than absolutes, she concedes that exceptional forms such as *'Kun.din.-nen* discussed above reveal the conflict between prosodic tendencies and morphological requirements; in these cases, morphological demands can take presedence over prosody, thereby upsetting the prosodic template.[6] Thus, in multimorphemic stems, the trochaic template can be outweighed

---

6   In Smith's account, the prosodic tendency would be outweighed by the demands of morphological suffixes such as *-in* and *-ung* to select specific plural suffixes.

by morphological concerns as the main stress remains on the stem syllable (e.g., 'Kö.ch-in, 'Kö.ch-in.-nen "female cook-s").

Building on this approach, Schuhmann and Putnam (2021) have also recently modeled the prosodic template for German plurals within a Distributed Morphology (DM) approach. There, a prosodic boundary for nominal plurals is established directly in the syntax through the generation of a Prosodic Word ω. Crucially, in this account, the prosodic unit for plural exponents is established only for a particular feature configuration in the syntactic structure (n[+pl]). Thus, in this first step, the syntax establishes just a prosodic boundary, devoid of any segmental information; as is typical in DM, the phonological realization of morphosyntactic features is achieved through the insertion of Vocabulary Items at Spell-Out (when morphosyntactic features are mapped onto phonological content). Similar criticisms as above can be raised here about how exceptions are modeled formally. While approaches proposed by Schuhmann and Putnam (2021) and Smith (2004, 2020) did not address this question, one option could include word-level-specific tags or features—perhaps inherent in the derivational suffixes discussed above—that would mark individual words such as multimorphemic stems as exceptional. This could also include marking some words differently which have been borrowed but not integrated into German. The feature configuration with this additional component (tag or feature) would no longer trigger the necessary prosodic boundary for trochaic plural exponents.

Other generative linguistic perspectives have analyzed the prosodic generalization for German noun plurals with Optimality Theory (OT, e.g., Prince, 1993/2004). In Wiese's (2009) account, the prosodic pattern was formalized as the markedness constraint "Trochee" interacting with other constraints. For this analysis, Wiese (2009) postulates constraints that are specific to plural or singular forms of the morphological paradigm. In particular, the central markedness constraint Trochee is specific to plural forms, requiring that they end in a trochee. Faithfulness constraints are likewise specific to certain parts of the morphological paradigm, i.e., singular vs. plural forms. Overall, at the core of this paradigmatic analysis, Trochee is ranked higher than a faithfulness constraint against segment insertion in plural forms (DEP-SegPl) but lower than a faithfulness constraint against segment insertion in singular forms (DEP-SegSg). Clearly, one central drawback of such an analysis is that the crucial constraints are assumed to be specific to either the plural forms or the singular forms of the morphological paradigm, thus reducing the explanatory factor in this model. Here, to achieve the prosodic pattern in plurals, very specific constraints have to be stipulated for specific parts of the paradigm. Since in OT, all constraints are universal and violable, the very specific formulation of constraints made to fit the German plural data makes them less promising as constraints of the universal set of OT constraints.

In another recent OT account of German plurals, Trommer (2021) argues that the prosodic shape of German plural forms arises as an Emergence-of-the-Unmarked effect (TETU, McCarthy and Prince, 1994). Trommer's model captures both the various allomorphs and the prosodic pattern of German plurals in a unified account in which he combines older phonological proposals with recent developments within OT. Specifically, his account is situated within Stratal OT (Kiparsky, 2015; Bermúdez-Otero, 2018) and Autosegmental Colored Containment Theory (Revithiadou,

2007; van Oostendorp, 2008; Zimmermann, 2014, 2017). Notably, Trommer (2021) does not assume a prosodic foot template for plurals, but argues that this pattern arises as a TETU effect (McCarthy and Prince, 1994). This is achieved in large part by resurrecting an older proposal that the representation of schwa-affixes constitutes a defective segment-sized phonological unit which, according to Trommer (2021, p. 605), corresponds to the C- or X-slot in Wiese (1986), Hall (1992), and Noske (1993). Given that faithfulness constraints protect incomplete segments less strictly, the templatic metrical effects emerge as a TETU phenomenon. By proposing that there is only one general underlying plural affix, the account also achieves an analysis in which the prosodic pattern holds for all plural forms. This single, general underlying plural affix consists of an underspecified root node and a floating coronal feature. If this general, phonologically underspecified, featureless affix is not combined with any other suffixes, Trommer suggests that the phonology will spell it out as schwa, the most unmarked segment in German. The general affix can also be combined with more restrictive plural affixes, either a feminine or a neuter plural marker, thereby specifying additional segmental information, such as a nasal or a pharyngeal feature. Finally, drawing on the pre-OT literature, Trommer (2021) proposes that umlaut is triggered by affixes containing a floating feature which conditions vowel fronting, here a floating coronal feature. Overall, Trommer (2021) consternates that, unlike Wunderlich (1999), most proposals cannot achieve a natural account of the prosodic regularity, the allomorphs, and the "implicational relationships between umlaut and suffixation" (p. 650).

Thus, prosody figures prominently in many—albeit not all—formal models of German plural formation, including in Köpcke and colleagues' schema-based accounts. Here, in particular, clinical studies have assigned a more central role to the trochaic patterns within schematic accounts of plurals when discussing a case study of a patient with aphasia (Domahs et al., 2017), and child language acquisition in impaired and non-impaired language development contexts (Kauschke et al., 2013), as discussed above. This work suggests that prosody is the constraining framework within which plural morphology is built. The PTH introduced above similarly proposes that L2 prosodic-phonological representations are the basis for inflectional morphology in L2 development. Thus, we set out to test whether and how these prosodic representations might develop in adult L2 acquisition in a foreign language learning context.

2.5 Filling the gap: testing prosodic sensitivity in German plurals

In this work, we focus on whether the prosodic pattern that has been established in descriptive and theoretical analyses of German plural forms can be considered a productive part of the grammatical system of L1 German users, and how it might develop in L1 English-L2 German users across language proficiency levels. This aspect of the study aims to address if and when L2 German users develop intuitions about a *potential* prosodic condition for German nominal plural formation. In other words, we ask whether this metrical pattern discussed for German plurals is merely the

result of historical developments in German (see Smith, 2022 for an overview), or whether it is now an existing but perhaps non-productive pattern in the German grammatical and/or lexical system. One way to test productivity is to assess how speakers deal with non-words. While the trochaic pattern occurs in many existing German words as well as in integrated loanwords such as *Ga.ˈra.ge-n* "garage-s", to our knowledge, empirical studies of healthy adult language users have not yet focused on the role of prosody in the production of novel plural forms or in well-formedness judgments by (healthy) language users and during adult L2 acquisition (but see Vogt, 2016 for a pilot study with L2 German learners).

To this end, we developed two tasks to assess the role of prosody in productive and receptive German plural formation. The first task (L1 and L2 German participants) was a plural elicitation task to examine the role of prosody in the production of plural forms in non-words. Just for the L2 German users, we also included a plural elicitation task with existing German words.[7] The third task (L1 and L2 German participants) was then a grammatical acceptability judgment task (GAJT), which examined the role of prosody in the ratings of existing and made-up plural formations of existing German words. Note that both English and German have been characterized as languages containing the prosodic constituent "foot". Similar to German, evidence supports the claim that English "builds binary trochaic feet" (Garcia and Goad, 2021, p. 20; see also Domahs et al., 2014), an aspect to which we will return in the discussion. Thus, our tasks were designed to answer the following research questions respectively:

- Task 1: *Plural elicitation task—non-words*

  - RQ1a: To what extent do L1 German and L1 English-L2 German speakers produce trochaic plurals for non-words?
  - RQ1b: Does L2 German speakers' use of trochaic plural forms for non-words increase with increasing proficiency?

- Task 2 (L2 users only): *Plural elicitation task—existing words*

  - RQ2a: How accurately do L1 English-L2 German speakers produce plural forms for existing German nouns? Does accuracy improve with increasing proficiency?
  - RQ2b: To what extent do L1 English-L2 German speakers produce trochaic plural forms for existing German nouns? Does the number of trochaic plurals increase with increasing proficiency?
  - RQ2c: To what extent do L1 English-L2 German speakers produce trochaic plurals forms for incorrectly produced existing German nouns?

- Task 3: *Plural well-formedness rating task*

  - RQ3a: To what degree do well-formedness judgments of L1 German and L1 English-L2 German users correlate with the plural forms' prosodic structure? I.e., do users rate plural

forms that are *non*-trochaic as less well-formed compared to forms that are *trochaic*, both within correct and within incorrect plurals?
  - RQ3b: Do L2 German users' ratings of the well-formedness of correct and incorrect (trochaic and non-trochaic) plural forms change with increasing proficiency? I.e., is the correlation between plural type (correct/incorrect by trochaic/non-trochaic) stronger in more proficient users?

The methodological details of the study tasks, the procedures, and the participants are described in the following sections.

3 Empirical study: probing prosodic intuitions of L1 and L2 users

3.1 Participants

As noted, data were collected from two major groups of participants: (1) L1 users of German and (2) L1 English-L2 German users. A total of 54 (21 male, 33 female) L1 German users participated, of which we will present the data from 30 (10 male, 20 female) participants below 40 years old (mean 26, range 18–38), which is closer to the age-range of the L2 participants (mean: 21.5, range 18–28). A total of 98 (48 male, 50 female) L1 English-L2 German users participated in the study (however, $n = 97$ for Task 3). All participants in this group were enrolled in German classes at a large private university in the U.S. Here, we use semester-level (plus immersion experience) as a proxy for proficiency. L2 participants were assigned to one of four proficiency group levels based on the number of semesters they had studied German. Twelve (m = 4, f = 8) students were in their 1st or 2nd semester German classes ("Year 1"), 15 (m = 2, f = 13) participants were in their 3rd or 4th semester German classes ("Year 2"); 10 (m = 5, f = 5) participants were enrolled in a 5th semester course or above and had spent less than 6 months in a German-speaking country ("Year 3"). Sixty-one participants (m = 37, f = 24) were enrolled in a 5th semester course or above and had at least 6 months of immersion experience ("Year 3 + Imm."). These participants comprised the largest group because of the tendency for German students at this university to have spent 16–22 months in a German-speaking country.

3.2 Methodology

3.2.1 Task 1: non-word plural elicitation task

The first task was a non-word plural production task, similar to a *wug*-test (Gleason, 1958). This task was intended to examine L1 and L2 German users' prosodic intuitions for plural formation by testing how frequently they would produce trochaic plural forms. The non-words were presented with an indefinite article (*ein/e* "a, masc. or neutr. / fem.") and participants were asked to provide a plural form for each word. To respond, participants were asked to both say out loud and type how they would form the plural for each given singular non-word. To help facilitate the plural formation, they were given a number to use for the plural production. For

---

example, participants saw and heard *ein Schliemo* ("a Schliemo") and responded *3 Schliemos* (or other plural options).

In total, participants responded to a total of 76 non-word tokens. Some non-words were drawn from other studies on German plurals (Hahn and Nakisa, 2000; Kauschke et al., 2013; Köpcke, 1998), while others were created by the authors. Non-words were created so as to respect German phonotactics and co-occurrence restrictions between suffixes and stem-vowel changes (Augst, 1979; Wurzel, 1998; *see above*). Non-words were distributed across seven categories, which are presented with example non-words for each category in Table 2.

Of these, all but one category, namely the tokens ending in an unstressed syllable with a full vowel, e.g., *Schliemo,* included non-words that were expected to form their plural by applying the prosodic principle. These vowel-final tokens were expected to form plurals using *-s*, which is an ending that does not have to conform to the prosodic principle (see Section 2.2 above); however, since the singular forms are trochaic, an *-s* suffix would still lead to a trochaic plural form.

Effectively, for monosyllabic and polysyllabic words with final stress, participants were expected to add a syllabic suffix (<-en> /-ən/, <-e> /-ə/, or <-er> /-ɐ/, respectively) to achieve a word-final trochee. For the remaining words (words in various schwa-final-syllables), participants were expected to add non-syllabic suffixes (<-n> /-n/, -Ø, or <-s> /-s/).

3.2.2 Task 2: real word plural elicitation task

This task was presented only to the L2 participants as a pre-task to examine their accuracy with and prosodic intuitions for plural formation for existing words in German. This task mirrored Task 1, the plural elicitation task with non-words, except that it was conducted with existing words from the German lexicon. Here, participants responded to a total of 151 real word stimuli, with 122 critical stimuli and 29 fillers. Fillers included words with suffixes, e.g., *'Lös-ung* "solution", or words with stressed prefix-like elements, e.g., *'Ab-schied* "farewell". The critical stimulus words were distributed across eight categories, which are presented with example stimuli in Table 3.[8] Notably, each participant responded to 106 critical stimuli where the prosodic principle was expected to take effect.

Just as in Task 1, the words in all but one category were chosen because each of them forms its plural by applying the prosodic principle. Thus, tokens ending in an unstressed full vowel, e.g., *Auto* "car", were included here as well. These forms are not guided by the prosodic principle but form plurals in *-s* in SG. Even when participants do not build accurate plural forms, this task allowed us to test whether L2 users apply the prosodic principle when forming plurals. Thus, to achieve a word-final trochee, participants would

need to add a syllabic suffix to form plurals for monosyllabic and polysyllabic words with final stress. For other words (words in open or closed schwa-final-syllables), participants were expected to add non-syllabic suffixes (including the null-suffix) to keep a word-final trochaic word form in the plural.

3.2.3 Task 3: grammatical acceptability judgment task (GAJT)

Task 3 was a grammatical acceptability judgment task (GAJT) to examine L1 and L2 users' intuitions of plural formation by testing whether the (lack of a) trochaic pattern contributes to the well-formedness judgments. To this end, participants were presented with various plural forms of real German words, including both correct and incorrect, i.e., non-standard German plural forms. For instance, the task included *der Schlüssel—die Schlüssel* (correct) "the key—the keys" as well as incorrect/non-standard *die Schlüssels, die Schlüsseln*, etc. Participants saw (and heard) the singular form with the definite article (*der*, *die*, *das*, "the, sing. masc., fem., neutr."), followed by the plural form twice, once with the definite plural article (*die*, "the, plural"; no gender distinctions), and once with *viele* "many", e.g., *der Schlüssel—die Schlüssels, viele Schlüssels* "the key—the keys, many keys".

After reading and hearing the singular and plural forms, participants were asked to repeat the plural form with *viele* "many", and then asked to rate the well-formedness of the provided plural forms on an 8-point Likert Scale, which was visualized with an arrow pointing in both directions. The left side of the arrow, numbers 1-4, were presented in green, while the right side, numbers 5-8, were presented in red. Thus, there was no mid-point and no neutral point on the scale; participants were asked: "Wie GUT oder SCHLECHT klingt diese Pluralform?" ("How GOOD or BAD does this plural sound?"), whereby *good* was presented in green and *bad* in red. The options on the Likert Scale were labeled with the phrases "Für mich klingt diese Pluralform____" ("For me, this plural sounds _____"). The values spanning 1-2 and 3-4 received the green labels "SEHR GUT" ("REALLY GOOD") and "EHER GUT" ("SOMEWHAT GOOD"), respectively, while 5-6 and 7-8 received the red labels "EHER SCHLECHT" ("SOMEWHAT BAD") and "SEHR SCHLECHT" ("REALLY BAD"), respectively.

This task included 179 tokens. This included the correct and various incorrect, i.e., non-standard plural forms for 42 existing lexemes of the German lexicon.[9] Crucially, incorrect forms included trochaic and non-trochaic forms. For example, for the stimulus *der Schlüssel* "the key", the correct plural form *die Schlüssel*, as well as the incorrect plural forms

---

8   *Steuer* occurred as feminine ("tax") and neuter ("helm, steering wheel"); these real words differ in their SG plural forms, *Steuern* and *Steuer*, respectively. The singular *Motor* "motor" occurred with two stress patterns, thus in two categories: final stress and stress shift. *Kaffee* "coffee" also occurred with two stress patterns (initial and final stress) in the singular, but was always part of the vowel-final category.

---

9   Note that this task was designed to focus on suffixes, and avoided, for the most part, testing plural marking by means of umlaut. Two of the stimuli form the plural with umlaut in SG (*Garten—Gärten* "garden—gardens", *Ofen—Öfen* "oven—ovens"), and all plural forms presented in the task included umlaut for both of these stimuli. Additionally, five nouns (*Mädel* "girl", *Löffel* "spoon", *Schlüssel* "key", *Schüssel* "bowl", *Tür* "door") had an umlaut in the orthography of the singular and the plural form in SG, which was kept for all plural forms presented in this task.

TABLE 2 Categories and examples of tokens used in Task 1 (*Plural elicitation task—non-words*).

Higher-level categories (singular form)	Categories (singular form)	Sub-categories (singular form)	Examples	Application of prosodic principle expected?
Final schwa-syllable (*n* = 26)	**Disyllabic, final schwa-syllable** (*n* = 26)	-*er* (*n* = 7)	*'Zolger, 'Treiker*	Yes
		-*el* (*n* = 7)	*'Spotel, 'Wenfel*	
		-*e* (*n* = 7)	*'Dalle, 'Jechte*	
		-*en* (*n* = 5)*	*'Tefen, 'Zaupen*	
No final schwa-syllable (*n* = 50)	**Monosyllabic** (*n* = 26)	NA	*'Kland, 'Jent*	
	Multisyllabic, no final schwa-syllable (*n* = 24)	Final stress (*n* = 19)	*Pintala'kor, Flasi'tar*	
		Final full vowel (*n* = 5)	*'Schliemo, 'Doscha*	No

Two words each from the -*er*, -*el*, and -*e* classes and eight from the monosyllables were presented with both *ein* "a, masc./neutr." and *eine* "a, fem." (*Words ending in -*en* are never feminine in SG.)

TABLE 3 Categories and examples of tokens used in Task 2 (*Plural elicitation task—existing words*).

Higher-level categories (singular form)	Categories (singular form)	Sub-categories (singular form)	Examples	Application of prosodic principle expected?
Final schwa-syllable (*n* = 58)	**Disyllabic, final schwa-syllable** (*n* = 58)	-*er* (*n* = 17)	*'Lehrer, 'Leiter*	Yes
		-*el* (*n* = 15)	*Ar'tikel, 'Kugel*	
		-*e* (*n* = 17)	*'Wolke, 'Name*	
		-*en* (*n* = 9)*	*'Kuchen, 'Wagen*	
No final schwa-syllable (*n* = 64)	**Monosyllabic** (*n* = 19)	NA	*'Berg, 'Lied*	
	Multisyllabic, no final schwa-syllable (*n* = 45)	Final stress (*n* = 24)	*Universi'tät, Stu'dent*	
		Nouns with stress shift (*n* = 5)	*'Autor, Di'rektor*	
		Final full vowel (*n* = 16)	*'Auto, 'Opa*	No

The stimuli also included all three genders (*der/die/das*: "the, masc./fem./neutr."), except for nouns with stress shift (all masculine). (*Words ending in -*en* are never feminine in SG.)

Schlüssele and *Schlüsseln* were presented. The categories used in this GAJT task, as well as example words for the lexical items and incorrect tokens, are shown in Table 4. These categories were meant to parallel those in the two plural elicitation tasks.

To summarize, this GAJT task was designed to test whether the prosodic shape, here trochaic vs. non-trochaic patterns within correct and incorrect plural forms, correlates with the well-formedness ratings of plural forms within the groups of L1 and L2 German participants.

3.3 Procedure

The three tasks were conducted using E-Prime (version 2.0). For each task, stimuli were presented one by one to participants visually on a laptop screen and auditorily via headphones at the same time. The auditory presentation of the stimuli was intended to ensure that all participants could hear the stress pattern. These tokens were recorded by a female speaker in Germany who grew up as a monolingual L1 German speaker. She pronounced the real words, incorrect plural forms, and non-words in a natural way and at a natural

speaking rate.[10] The L2 participants began with Task 2, the plural elicitation task with real words. They then followed the same procedure as the L1 participants, completing Task 1 (plural elicitation task with non-words) followed by Task 3 (GAJT). At the end, each participant also completed a brief language background questionnaire.

3.4 Analyses

The research questions for all three tasks focus on the prosodic patterns in the elicited and provided plural forms. To this effect, the analyses and results presented here primarily focus on prosody, as well as a potential progression within the levels of L2 German users. For the non-word plural elicitation task (Task 1), results focus on whether the plural forms produced show a word-final trochaic form, which is analyzed as a binary categorical variable (trochaic vs. non-trochaic). For the plural elicitation task with

---

10 The speaker was instructed by author KS (an L1 German speaker) to pronounce all words as if they were real words. KS examined the recordings and chose those that sounded most natural in her judgment.

TABLE 4 Categories of word types used in Task 3 (*GAJT: grammatical acceptability judgment task*).

Higher-level categories (singular form; *n* lex. items; *n* tokens)	Categories (singular form; *n* lex. items; *n* tokens)	Sub-categories (singular form; *n* lex. items; *n* tokens)	Example lexical items	Example non-standard plural forms (tokens)
Final schwa-syllable (*n* = 20; 68)	**Disyllabic, final schwa-syllable** (*n* = 20; 68)	*-er* (*n* = 6; 24)	*'Fenster, 'Schwester*	*Fenstern, Fenstere*
		-el (*n* = 10; 32)	*'Löffel, 'Gabel*	*Löffeln, Löffelen*
		-en (*n* = 4; 12)	*'Kissen, 'Garten*	*Kissene, Kissenen*
No final schwa-syllable (*n* = 21; 67)	**Monosyllabic** (*n* = 13; 42)	NA	*'Fisch, 'Frau*	*Fisch, Fischen*
	Multisyllabic, no final schwa-syllable (*n* = 8; 25)	Final stress (*n* = 3; 11)	*Klau'sur, Sa'lat*	*Klausur, Klausure*
		Nouns with stress shift (*n* = 3; 9)	*Pro'fessor, 'Doktor*	*Professore, Professorn*
		Final full vowel (*n* = 2; 5)	*'Kino, 'Auto*	*Kino, Kinoer*

TABLE 5 Model summary for Percent Trochaic plural forms in the plural elicitation task with non-words in L2 and L1 speakers (Task 1).

	Estimate	Std. error	z-value	*p*-value
(Intercept)	7.87	1.10	7.15	<0.001
Level1: Y1 vs. Y2	0.30	0.43	0.69	0.488
Level2: Y1+2 vs. Y3	1.59	0.43	3.74	<0.001
Level3: Y1-3 vs. Y3+	0.50	0.29	1.70	0.089
Level4: L2 vs. L1	0.00	0.36	0.01	0.990
Class: *-el* vs. *-e*	−3.41	1.14	−3.00	0.003
Class: *-en* vs. *-e*	−3.80	1.14	−3.33	0.001
Class: *-er* vs. *-e*	−3.73	1.13	−3.30	0.001
Class: FinStress vs. *-e*	−4.69	1.11	−4.22	<0.001
Class: Mono vs. *-e*	−4.63	1.11	−4.17	<0.001
Class: -V vs. *-e*	−2.82	1.15	−2.47	0.014
Ending: *-e* vs. *-n*	0.13	0.20	0.67	0.504
Ending: *-en* vs. *-n*	−0.21	0.18	−1.12	0.261
Ending: *-er* vs. *-n*	−0.24	0.31	−0.76	0.445
Ending: -Ø vs. *-n*	−1.45	0.19	−7.70	<0.001
Ending: other vs. *-n*	−2.26	0.44	−5.15	<0.001
Ending: *-s* vs. *-n*	−2.55	0.24	−10.51	<0.001

real words (Task 2) completed only by the L2 users, results focus on (i) accuracy scores—a binary categorical variable (correct vs. incorrect), and (ii) just as in Task 1, whether the plural forms produced show a word-final trochaic form—a binary categorical variable (trochaic vs. non-trochaic). For Task 3, the grammatical acceptability judgment task, the data obtained by the experiment yielded ordinal ratings on an 8-point Likert scale. For the results presented here, we set aside questions of stem-vowel changes (umlaut) (but see Schuhmann and Smith, accepted).

3.4.1 Coding
3.4.1.1 Trochaic vs. non-trochaic
For both the non-word and real word plural elicitation tasks, we analyzed how often participants produced trochaic plural forms for the singular forms provided. A form was coded as "trochaic" when the plural form ended in a sequence of a stressed followed by an unstressed syllable. All other forms were coded as "non-trochaic", i.e., most notably, monosyllabic forms, forms with a final stressed syllable, or a sequence of two final reduced syllables. All items were included in the analyses.

3.4.1.2 Ending chosen
When participants produced a form in either of the plural elicitation tasks that did not differ from the singular form provided, this null-marking was coded as "NO". Thus, we conceptualized this as a zero or null-morpheme (a morpheme with no overt phonetic content). When an ending did not correspond to any of the typical plural markers for SG, i.e., *-e, -en, -n, -er, -s,* or *-Ø,* it

was coded as "other" (e.g., <-in> or deleting final consonants, e.g., turning *Rücken* "back" into *Rücke* "backs"). Note that these are the suffixes used by the speakers in the study, so the occurrences of individual suffixes are not evenly distributed; in particular, both *-er* and "other" occurred rarely as suffix-markers. Umlaut is not considered in this analysis of suffixes chosen.

3.4.1.3 Accuracy

For the real word task, we analyzed accuracy, i.e., whether the plural forms produced by the L2 users corresponded to "correct", i.e., target-like plural forms in terms of the overall *segmental* composition of the plurals. A form was coded as correct if it corresponded to SG or frequently used forms in colloquial German, such as *Mädel* "girls" or *Mädel-s* "girls, colloq.", respectively, which are also listed in Duden-Deutsches Universalwörterbuch (2023).

3.4.2 Statistical analyses of trochaic pattern and accuracy in Tasks 1 and 2

For the binary categorical variables (trochaic vs. non-trochaic, correct vs. incorrect), the data were analyzed with a binomial generalized linear mixed-effects model in R (version 4.2.2; R Core Team, 2021) using the lme4 package (version 1.1-31) (Bates et al., 2015). We included either "Percent Trochaic" (trochaic vs. non-trochaic) or "Percent Accurate" (accurate vs. inaccurate) as the binary dependent variable. In a generalized linear mixed-effects model, the binary dependent variable is transformed into a continuous variable via a log odds link function, logit(p), i.e., the natural log of the probability. Probability here refers to the probability of providing a trochaic answer.

We included proficiency level ("Level") as a fixed effect in the model. Level had four levels of increasing German language proficiency for the real word task with just L2 users (Year 1, Year 2, Year 3, Year 3 + Imm.), and five levels of increasing German language proficiency (Year 1, Year 2, Year 3, Year 3 + Imm., L1) for the non-word task with L1 and L2 participants. This factor was reverse-Helmert-coded, where each level of a factor is compared to the mean of the previous levels. Thus, in the model outputs, Level1 represents the contrast between the mean of Year 1 and Year 2, Level2 the contrast between the mean of Years 1 and 2 vs. the mean of Year 3, Level3 the contrast between the mean of Years 1, 2, and 3 versus the mean of Year 3 + Imm., and—for the non-word task additionally—Level4 the contrast of the mean of all L2 levels (Year 1, Year 2, Year 3, Year 3 + Imm.) vs. the mean of L1.

We fitted models with by-subject and by-word random intercepts and a random slope for Level by word. Since a stimulus' membership in a word class ("Class") and—at a reviewer's suggestion—the chosen ending ("EndingChosen") could have an effect on the prosodic shape, we tested whether Class and EndingChosen should be included in the model as covariates for Percent Trochaic by means of model comparisons (cf. Baayen et al., 2008).[11] Class refers to the stimulus category of a singular word (labeled "categories" and "sub-categories" in Tables 2–4)

based on features such as its prosody, stress pattern, and word-final phonology. The seven levels of Class were: (singular words ending in) *-e, -el, -en, -er*, FinStress (words ending in a stressed syllable), Mono (monosyllabic words), and -V (words ending in an unstressed full vowel). Task 2 with real words additionally included the Class category of StrChg (words with a stress change between singular and plural forms). EndingChosen refers to the specific plural suffixes that participants provided when pluralizing each of the singular stimulus forms. The seven levels of EndingChosen were: the suffixes *-e, -n, -en, -er*, NO (null morpheme/no phonetically overt suffix), *-s*, and other. Class and EndingChosen were each treatment-coded, which compares each level of Class or EndingChosen to a reference level. For Class and EndingChosen, we selected the most predictable forms as the reference level. Thus, for Class, we chose singular forms ending in *-e*, for which the plural form is unambiguously the *-n* suffix (see also Kauschke et al., 2013). Conversely, for EndingChosen, the reference level was the *-n* suffix, as this is the predictable plural suffix for a subset of stimuli, namely singulars ending in *-e* (note, however, that the *-n* suffix also occurs on feminine singulars ending in a closed schwa-syllable).[12] *Post-hoc* pairwise contrasts for Level were calculated with the *emmeans*-function (Lenth, 2021). The statistical analysis for the Likert scale-like data in the grammatical acceptability judgment task (GAJT) in Task 3 is discussed in the relevant section (Section 4.3) below.

4 Results

4.1 Results Task 1: non-word plural elicitation task

4.1.1 Percent Trochaic

We first analyzed how often participants produced trochaic plural forms in the non-word task. Figure 1 presents the mean Percent Trochaic values and 95% confidence intervals (CI) for participants' plural forms at each (proficiency) Level (L2: Year 1, Year 2, Year 3, Year 3 + Imm., and L1). With 0.92 (SD 0.27), the vast majority of plural forms produced by L1 German users were trochaic. The L2 users produced an increasing number of trochaic plural forms for non-words with increasing L2 experience. Their mean percentage scores ranged from 0.78 (SD 0.41), to 0.87 (SD 0.34), 0.94 (SD 0.24), and 0.95 (SD 0.23) trochaic plural forms in the first year, second year, third year, and third year plus immersion, respectively. Of note is also that the means for L2 users in the two Year 3 groups (Year 3 and Year 3 + Imm.) showed a slightly higher number of trochaic plural forms than the L1 German users in this task.

Model comparisons within a binomial generalized linear mixed-effects model analysis and "Percent Trochaic" (trochaic vs. non-trochaic) as the binary dependent variable revealed that model fit improved when including (word) Class as a covariate

---

11   A model with an interaction of Level, Class, and EndingChosen would be too complex, due to the large number of factor levels for Level ($n = 5$), Class ($n = 7$; for real words: $n = 8$), and EndingChosen ($n = 7$), as well as imbalances in the data at hand.

12   Thus, the Class and EndingChosen comparisons were as follows. For Class: "Class: *-el* vs. *-e*", "Class: *-en* vs. *-e*", "Class: *-er* vs. *-e*", "Class: FinStress vs. *-e*", "Class: Mono vs. *-e*", "Class: *-V* vs. *-e*", and "Class: StrChg vs. *-e*" (only included in Task 2 with real words). For EndingChosen: "Ending: *-e* vs. *-n*", "Ending: *-en* vs. *-n*", "Ending: *-er* vs. *-n*", "Ending: NO vs. *-n*", "Ending: *-s* vs. *-n*", and "Ending: other vs. *-n*".
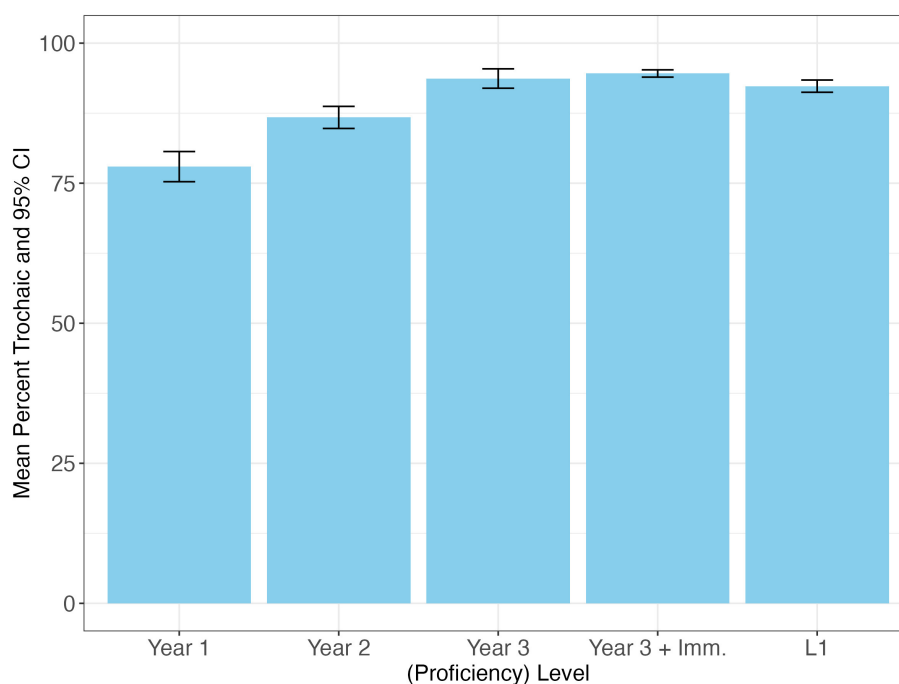
FIGURE 1
Percent Trochaic plural forms in the plural elicitation task with non-words (Task 1) in L2 and L1 speakers, split by (proficiency) Level.

($\chi^2_{(6)} = 60.5$, $p < 0.001$), and further improved when including EndingChosen ($\chi^2_{(6)} = 233.7$, $p < 0.001$). The model output (see Table 5) indicates that Level as well as Class and EndingChosen are significant predictors, as one Level contrast, as well as all of the Class contrasts and some EndingChosen contrasts were significant. *Post-hoc* pairwise comparisons for Level showed that the percentage of trochaic plural forms was significantly higher for more proficient L2-levels for the comparisons between Year 1 and Year 3 ($p < 0.005$) and Year 2 and Year 3 ($p = 0.012$), and marginal for the comparison between Year 1 and Year 3 + Imm. ($p = 0.053$), as well as Year 2 and Year 3 + Imm. ($p = 0.068$). None of the contrasts between individual L2 levels and the L1 level were significant.

These results suggest that overall, the percentage of trochaic forms in the production of plural forms for German non-words increased with increasing proficiency among L2 users, with significant differences between Year 1 and Year 3 (without immersion), and between Year 2 and Year 3 (without immersion). However, L2 users produced a similar number of trochaic plural forms as L1 users.

4.1.2 Percent Trochaic by (word) Class

The model summary from Table 5 above also indicates that (word) Class had an influence on Percent Trochaic. Due to the larger number of (word) Classes ($n = 7$) and (proficiency) Levels ($n = 5$), a Class-by-Level interaction was not included in the model.[13] Instead, we present visualizations for each (proficiency)
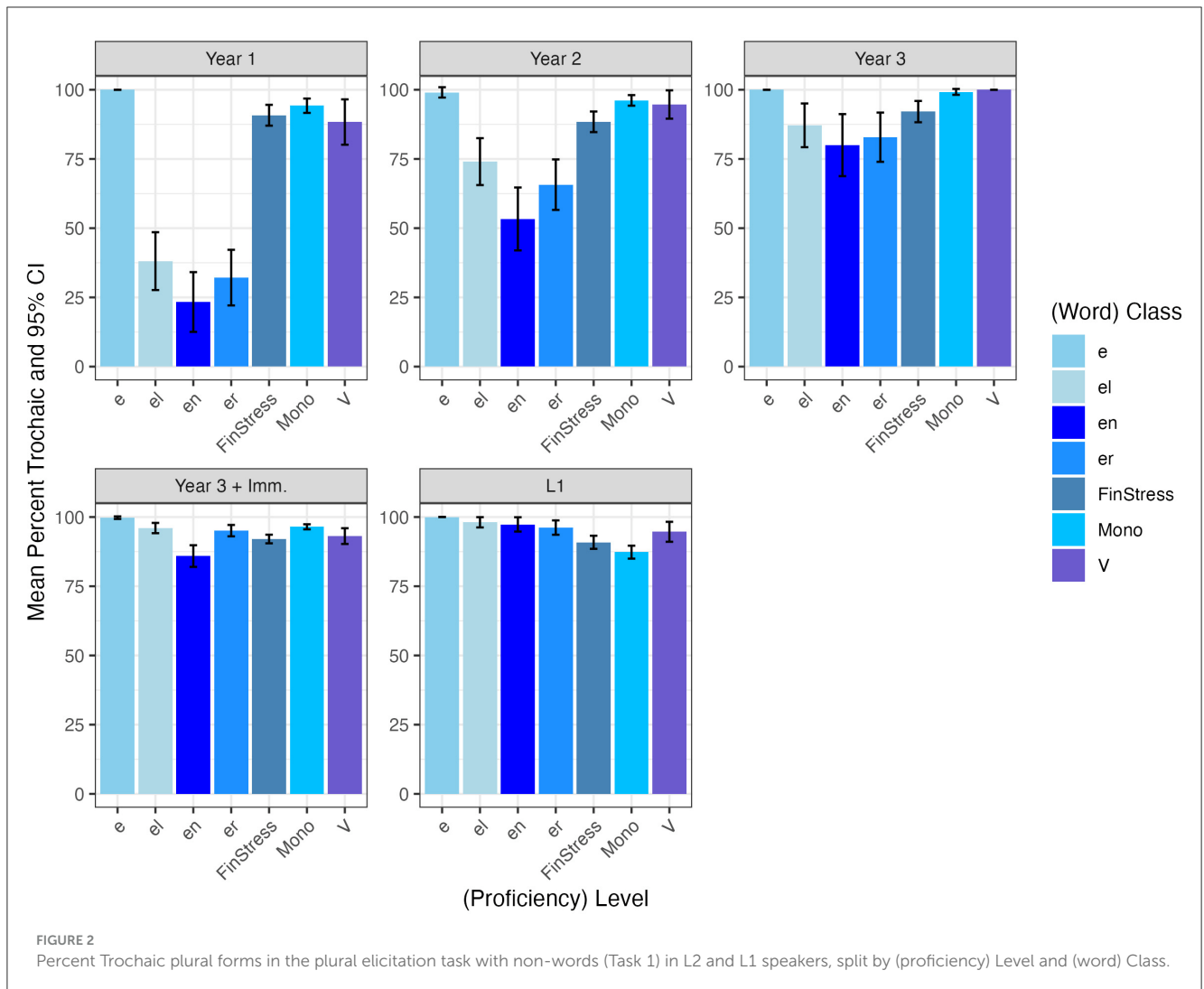
Level, which illustrate whether the percentage of trochaic plurals varied with (word) Class. The results are shown in Figure 2.

Overall, Figure 2 illustrates the effect of Class on the percentage of trochaic plural forms in non-words across (proficiency) Levels. The reference level of Class, singular non-words ending in *-e* (schwa), is listed first in each panel. For all groups, this level showed the highest percentage of trochaic plural forms, with close to 100% trochaic forms in each case. This can be accounted for by the fact that words in *-e* unambiguously form the plural with *-n* in German (cf. Kauschke et al., 2013). All comparisons between non-words in *-e* and the other individual Class levels are significant in the model summary (in Table 5) above, which collapses across all (proficiency) Levels. After words ending in schwa, the classes of words with final stress ("FinStress"), monosyllabic singulars ("Mono"), and words ending in an unstressed full vowel in the final syllable ("V") showed the next highest levels of trochaic plural forms at Years 1 and 2 (and to some degree at Year 3). L1 speakers also showed slightly lower percentages of trochaic forms for these three classes compared to words ending in schwa.[14]

Crucially, the graph further indicates that the lower percentage of trochaic plural forms in less proficient L2 German users was driven by the lower number of trochees in one specific group of non-words, namely non-words ending in closed schwa-syllables (*-el*, *-en*, *-er*). In other words, L2 users in Year 1, who showed the lowest percentage of trochaic plural responses in the analyses

---

13 This also means that *post-hoc* contrasts between Classes at individual Levels could not be conducted.

14 In fact, for the L1 speakers, the (word) Class of "Mono" (monosyllabic singulars) numerically showed the lowest amount of trochaic forms (0.87), followed by singular forms with final stress ("FinStress": 0.91), and then singular forms ending in a syllable with an unstressed full vowel ("V": 0.95).

FIGURE 2
Percent Trochaic plural forms in the plural elicitation task with non-words (Task 1) in L2 and L1 speakers, split by (proficiency) Level and (word) Class.

above (see also Figure 1), produced overwhelmingly trochaic plural forms in all classes except for the three classes of singular forms ending in pseudo-suffixes, i.e., in closed schwa-syllables (-el, -en, -er). The panels across the four L2 (proficiency) Levels (the first four panels in Figure 2) further indicate that the gap between the lower percentage of trochaic plural forms in words ending in closed schwa-syllables and words in an open schwa-syllable gradually closed over the course of the L2 developmental sequence. In the group of L1 users (the last panel), the three classes ending in pseudo-suffixes were all produced highly trochaically (0.96–0.98).

4.1.3 Percent Trochaic by EndingChosen

The model summary in Table 5 above also indicates that EndingChosen had an influence on Percent Trochaic. Thus, we asked whether and which specific suffixes were used more often with trochaic plural forms than others. The results are shown in Figure 3.

Figure 3 shows that at Year 1, L2 speakers produced the greatest number of trochaic plural forms using -n, followed by trochaic forms with the suffixes -e, -s, and -er. Fewer trochaic forms were produced with the -en suffix, the null-suffix ("NO"), as well as "other" (non-standard) suffix choices. By Year 2, both -en and
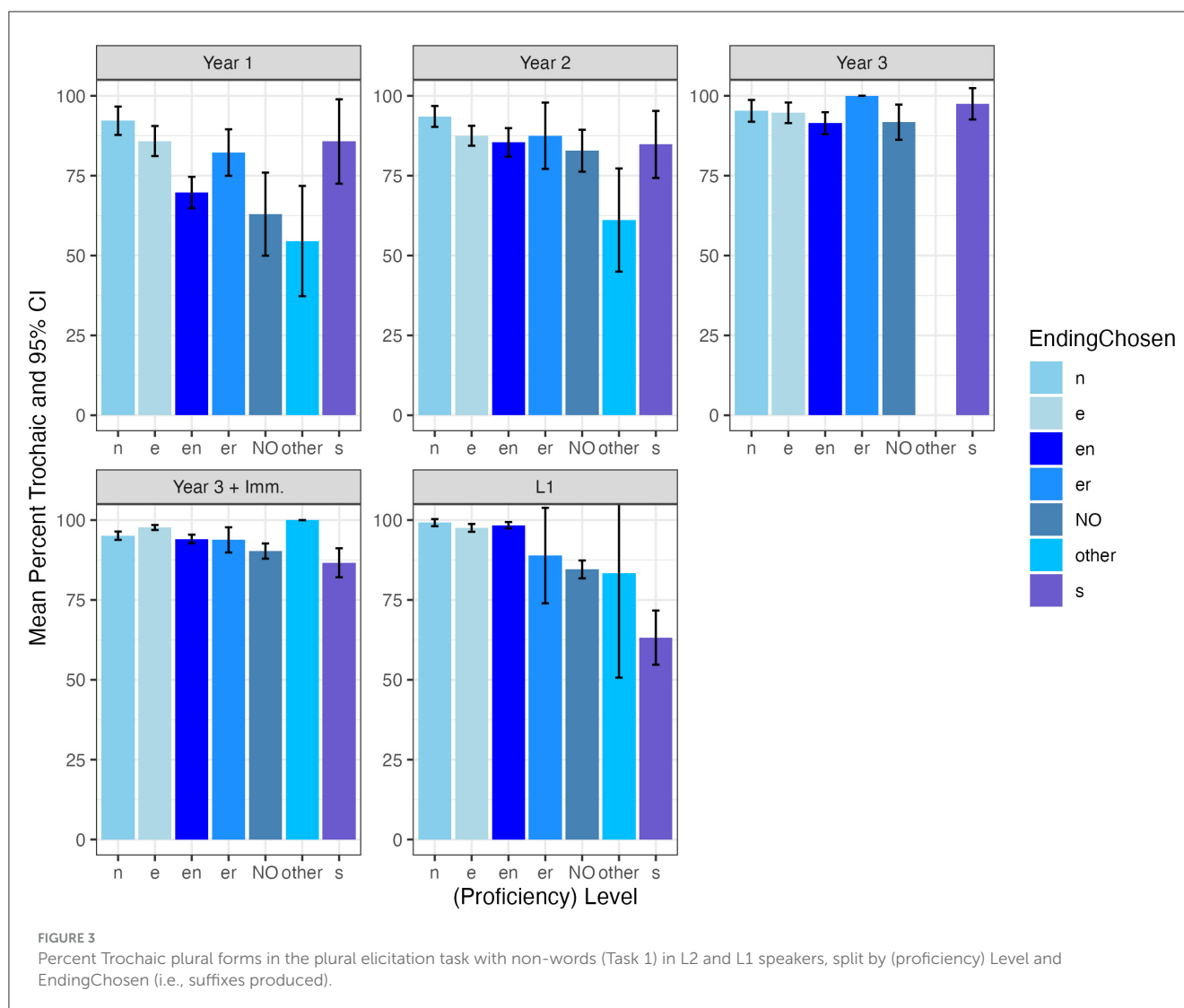
the null-suffix options were used to produce a similar number of trochaic plural forms as with the other existing suffixes (-n, -e, -s, -er). At Year 3 with and without immersion, L2 speakers showed a slightly higher percentage of trochaic forms for the suffixes.[15]

The L1 speakers showed trochaic plural forms with little variability when using the suffixes -e, -en, and -n, but fewer trochaic forms with the null-suffix, and even fewer trochaic plurals with the -s suffix (albeit with more variability). The suffix -er occurred very rarely in L1 users (n = 447, only 3.74% of the data), which is in line with the fact that this suffix has been frequently discussed in the literature as a non-productive plural allomorph. Both the -er and "other" endings (n = 138, only 1.15% of the data) were used with relatively trochaic plural forms but show large variability due to their very small number of occurrences.

4.1.4 EndingChosen based on (word) Class—across (proficiency) Levels

Figure 4 shows which suffixes participants chose for each (word) Class, broken down by (proficiency) Level. This graph

---

15 Except that the -s suffix (and possibly the null-suffix) occurred with slightly fewer trochaic plurals in Year 3 + Imm.

FIGURE 3
Percent Trochaic plural forms in the plural elicitation task with non-words (Task 1) in L2 and L1 speakers, split by (proficiency) Level and EndingChosen (i.e., suffixes produced).

serves to illustrate the combined information from the previous two figures. First, most of the differences between (proficiency) Levels can be seen for words ending in one of the three closed schwa-syllables. In these three (word) Classes, the choice of endings became more target-like (i.e., more similar to L1 speaker responses) with increasing proficiency. Combined with the discussion from the two previous sections, it is apparent that at Year 1, L2 learners used a fair number of *-en* suffixes—and some *-e* suffixes— for non-words ending in closed schwa-syllables, thus leading to trisyllabic forms, i.e., non-trochaic forms. This explains the substantial decrease in trochaic plural forms in these three (word) Classes at Year 1 in Figure 2, and the decrease of trochaic forms in comparison to other groups when the suffix *-en* was used at Year 1 in Figure 3.

Non-words ending in *-el* and *-er* also received more *-n* suffixes and fewer null-suffixes[16] at all L2 (proficiency) Levels compared to L1 speakers. At Year 2, the L2 users showed an increase in null-suffixes for all three closed schwa-syllables. At the same time, L2

users reduced their use of *-en* suffixes while increasing their use of *-n* suffixes for *-el* and *-er* classes from Year 1 to Year 2, resulting in more trochaic plural forms.[17]
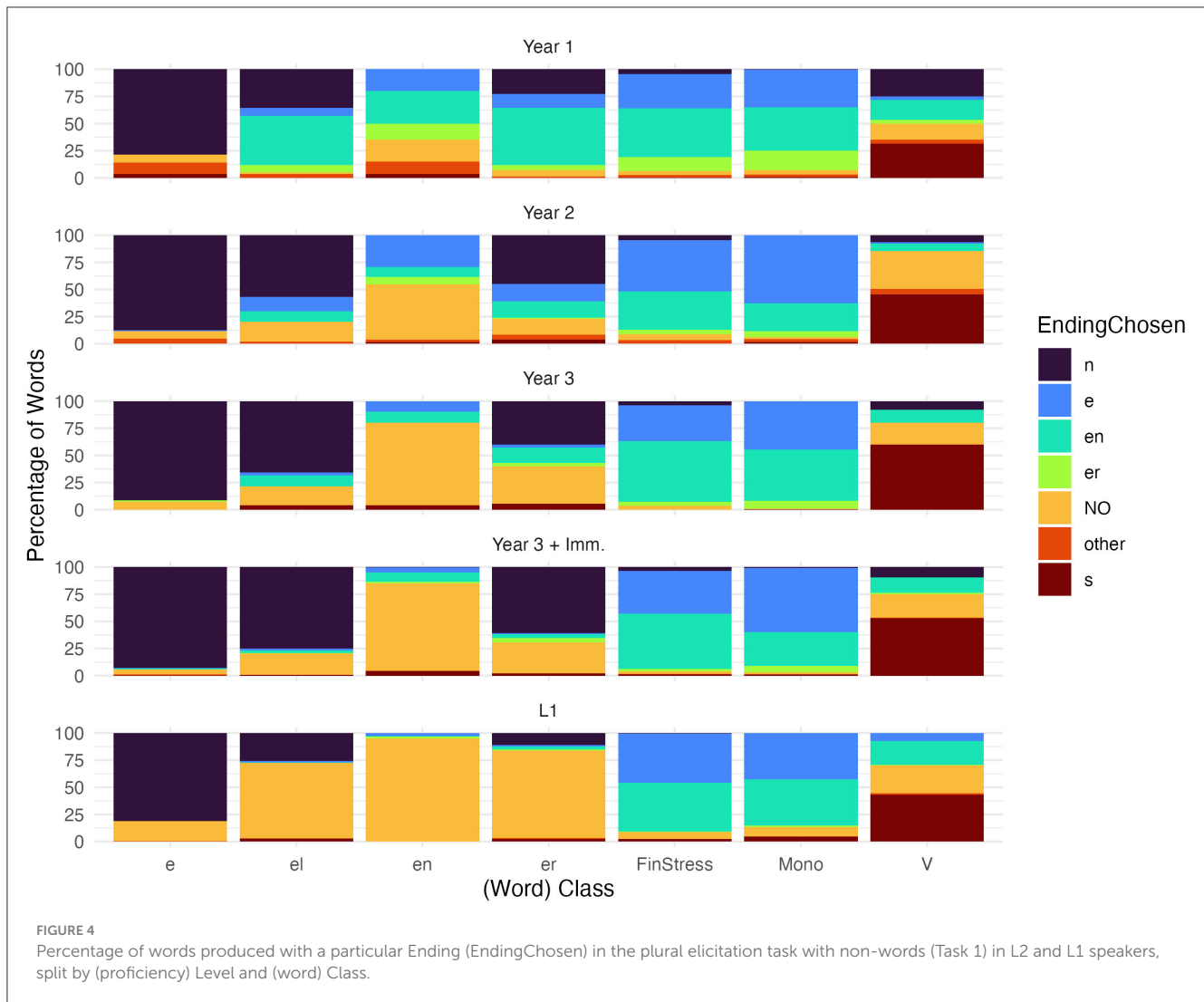
4.2 Results Task 2: real word plural elicitation task

4.2.1 Accuracy

For the real word task, which was completed by L2 speakers only, we first present the accuracy results. In Figure 5, the darker, left-hand bar within each (proficiency) Level represents the mean accuracy and 95% confidence interval for participants producing correct plural forms. As can be seen in this figure, L2 learners produced more correct plural forms with increasing L2 experience. L2 participants produced 0.41 (SD 0.49), 0.52 (SD 0.50), 0.66 (SD 0.47), and 0.66 (SD 0.47) correct plural forms in the first year

---

16   Non-words ending in *-en* received a fair number of *-en* suffixes, rather than *-n* suffixes, as the latter cannot be added to *-en* for phonotactic reasons.

17   There was also a small increase in *-e* suffixes for these three (word) Classes, in particular *-en* words, as well as singular words with final stress ("FinStress") and monosyllables ("Mono") by Year 2.

FIGURE 4
Percentage of words produced with a particular Ending (EndingChosen) in the plural elicitation task with non-words (Task 1) in L2 and L1 speakers, split by (proficiency) Level and (word) Class.

(Year 1), second year (Year 2), third year (Year 3), and third year plus immersion (Year 3 + Imm.), respectively.

Model comparisons within a binomial generalized linear mixed-effects model analysis and "Percent Accurate" (accurate vs. inaccurate) as the binary dependent variable revealed that model fit improved when including (word) Class as a covariate ($\chi^2_{[7]} = 88.6, p < 0.001$), and further when including EndingChosen as a covariate ($\chi^2_{[6]} = 386.1, p < 0.001$). The model output in Table 6 indicates that Level, Class, and EndingChosen were each significant predictors, as all Level contrasts were significant, as well as all Class and most EndingChosen contrasts. *Post-hoc* pairwise comparisons for Level showed that accuracy was always significantly higher for higher Levels ($p < 0.05$), with one exception: The contrast for Year 3 with and without immersion was not significant.

Overall, these results suggest that accuracy in the plural production of existing German words increased with course level, although the two Levels at the high end of the proficiency spectrum (Year 3 vs. Year 3 + Imm.) did not show significant differences between each other. At the same time, accuracy was far from ceiling for even the most advanced groups in our study, highlighting the difficulty in producing accurate plural

forms for L2 speakers. (Figures for accuracy split by Class and EndingChosen at the different (proficiency) Levels are available as Supplementary Figures 1, 2.)

4.2.2 Percent Trochaic

We next analyzed how often participants produced plural forms that were trochaic. In Figure 5 below, the lighter, right-hand bars present the mean Percent Trochaic values and 95% confidence intervals for participants' plural forms at each (proficiency) Level. As can be seen in this figure, the L2 learners produced a larger number of trochaic plural forms with increasing L2 experience. Their mean percentage scores ranged from 0.73 (SD 0.44), to 0.82 (SD 0.38), 0.90 (SD 0.30), and 0.92 (SD 0.27) trochaic plural forms in the first year, second year, third year, and third year plus immersion, respectively.

Model comparisons within a binomial generalized linear mixed-effects model analysis and "Percent Trochaic" (trochaic vs. non-trochaic) as the binary dependent variable revealed that model fit improved when including (word) Class as a covariate, ($\chi^2_{[7]} = 95.6, p < 0.001$), and further when including EndingChosen
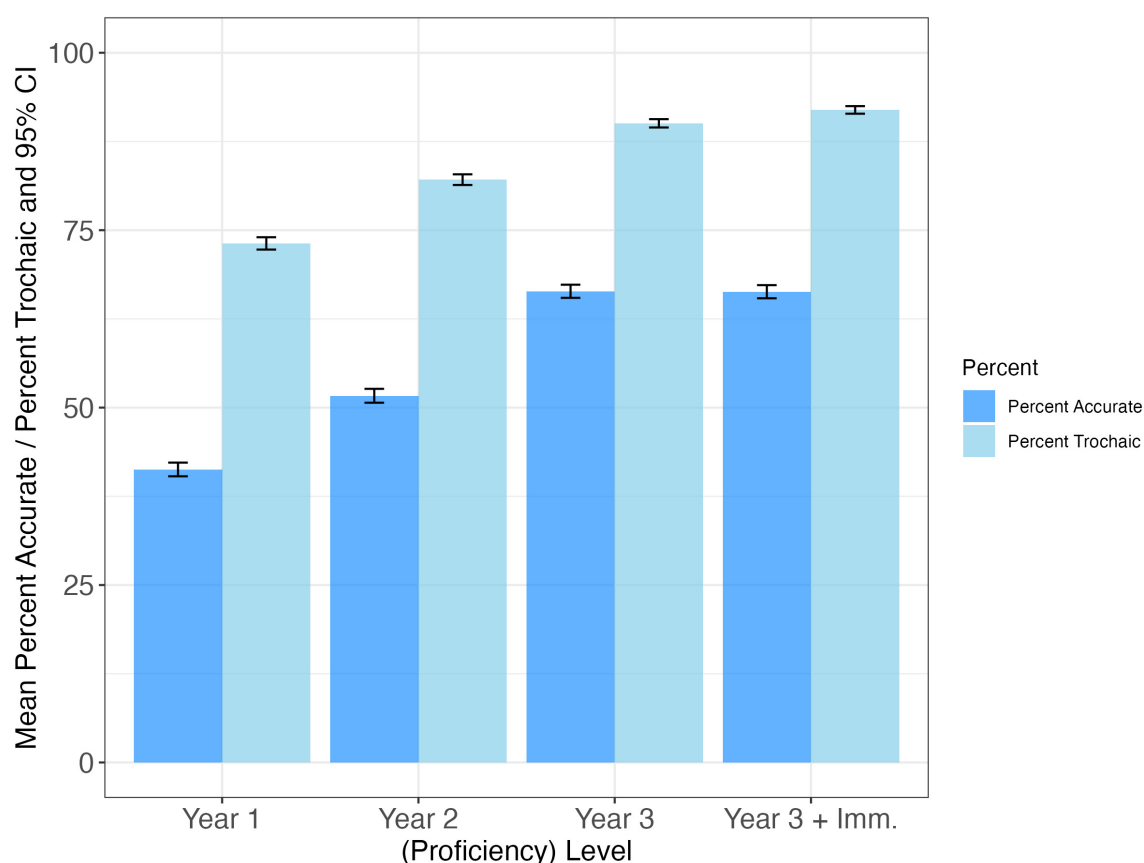
FIGURE 5
Percent Accurate (darker, left-hand bars) and Percent Trochaic (lighter, right-hand bars) plural forms in the plural elicitation task with real words (Task 2) in L2 speakers, split by (proficiency) Level.

($\chi^2_{[6]}$ = 215.3, p < 0.001). The model output (see Table 7) indicates that Level, Class, and EndingChosen were significant predictors, as all three Level contrasts were significant, as well as many of the Class and all EndingChosen contrasts. *Post-hoc* pairwise comparisons for Level showed that the percentage of trochaic plural forms was always significantly higher for higher Levels (p < 0.001), with two exceptions: The contrast between Year 1 and Year 2 was only marginally significant (p = 0.057), and the contrast for Year 3 with and without immersion was not significant.

Overall, these results suggest that the percentage of trochaic forms in the production of plural forms for existing German words increased with course level, although the low and high end of the proficiency spectrum (Years 1–2 and Year 3–Year 3 + Imm.) did not show significant differences from each other.

4.2.3 EndingChosen by (word) Class—split by (proficiency) Level

Figure 6 illustrates the effect of EndingChosen by (word) Class at the different (proficiency) Levels. This figure illustrates the combined information from Supplementary Figures 3 and 4, which show Percent Trochaic split by Class and EndingChosen, respectively, at the different (proficiency) Levels. Paralleling the non-word task, most of the development across (proficiency) Levels

took place in the Classes of words ending in closed schwa-syllables. First, all of these three Classes showed a substantial number of -*en* suffixes at Year 1, which were reduced in number by Year 2 and further reduced in number by Year 3 (without and with immersion). A small number of -*e* suffixes also appeared at Years 1 and 2. In tandem, the number of null-suffixes increased over the course of the four (proficiency) Levels, in particular for words ending in -*en* (and similarly for -*er*; yet Year 3 + Imm. had fewer null-suffixes for words ending in -*er* than Year 3). When comparing the patterns with those in the non-word task, L2 speakers at Year 1 started out using more null-suffixes for real words ending in -*el, -en, -er*, especially for words ending in -*en*.

4.2.4 Percent Trochaic of incorrect plurals

Further analyses of just the incorrect plural forms revealed that the overall grand mean percent of trochaic items decreased from 0.88 (accurate and inaccurate plurals) to 0.74 (only inaccurate plurals). When examining the percentage of trochaic forms among incorrect plurals across the (proficiency) Levels, an increase in prosodically well-formed—albeit incorrect—plural forms can be seen. In other words, when analyzing only the incorrect plural forms, the percentage of trochaic plural forms increased with increasing proficiency. Specifically, the L2 users across the four (proficiency) Levels produced incorrect but trochaic plural forms

TABLE 6 Model summary for Percent Accurate plural forms in the plural elicitation task with real words in L2 speakers (Task 2).

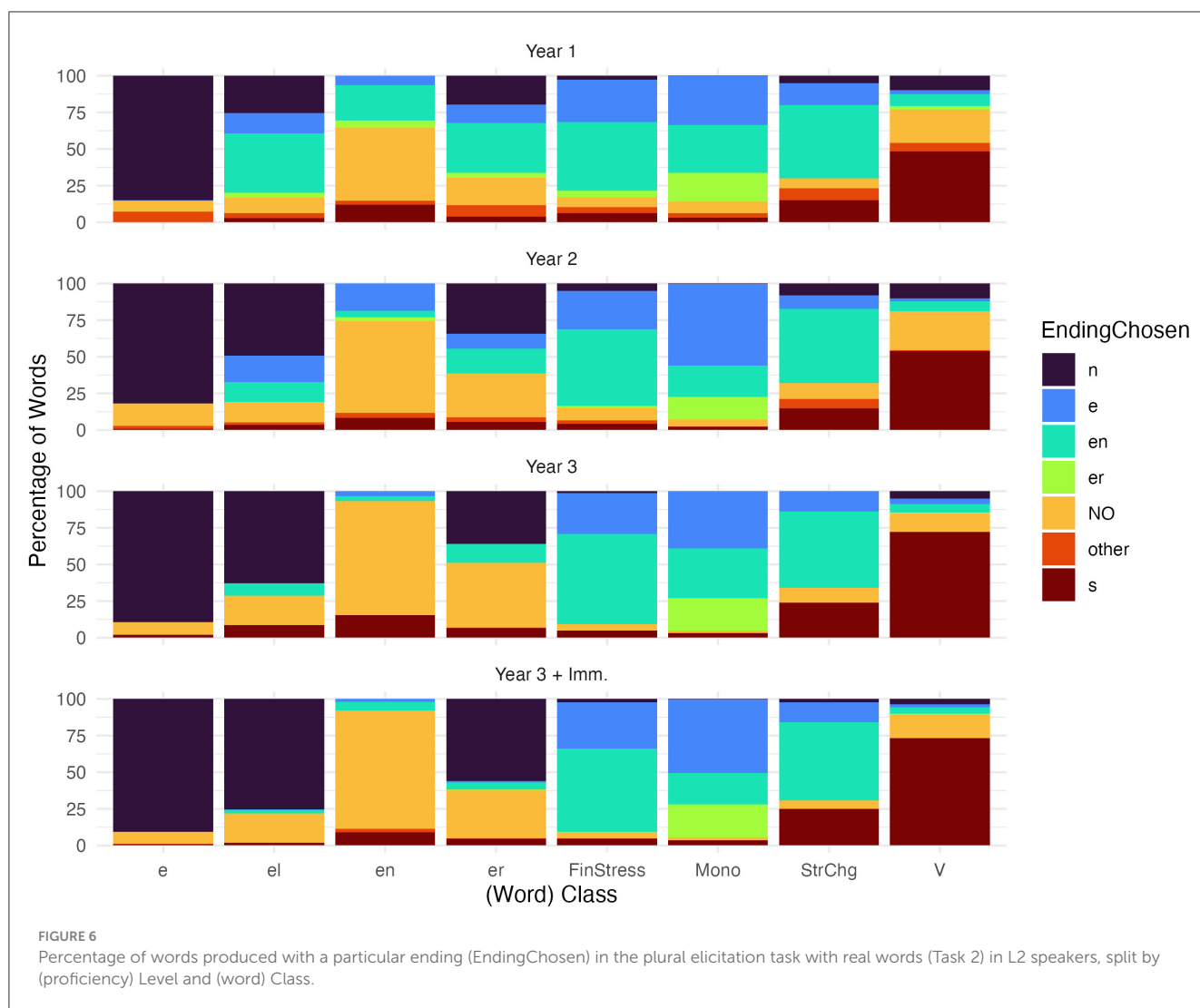|	Estimate	Std. error	z-value	p-value
(Intercept)	2.62	0.29	9.00	<0.001
Level1: Y1 vs. Y2	0.59	0.23	2.62	0.009
Level2: Y1+2 vs. Y3	1.22	0.22	5.58	<0.001
Level3: Y1-3 vs. Y3+	0.90	0.14	6.59	<0.001
Class: -el vs. -e	−4.73	0.41	−11.42	<0.001
Class: -en vs. -e	−3.21	0.46	−6.99	<0.001
Class: -er vs. -e	−3.87	0.39	−9.83	<0.001
Class: FinStress vs. -e	−2.82	0.37	−7.69	<0.001
Class: Mono vs. -e	−3.27	0.39	−8.43	<0.001
Class: StrChg vs. -e	−2.69	0.56	−4.83	<0.001
Class: -V vs. -e	−2.05	0.40	−5.17	<0.001
Ending: -e vs. -n	1.15	0.12	9.54	<0.001
Ending: -en vs. -n	0.89	0.11	8.08	<0.001
Ending: -er vs. -n	2.31	0.21	10.84	<0.001
Ending: -Ø vs. -n	0.75	0.09	8.27	<0.001
Ending: other vs. -n	−30.01	223,518	0.00	1.00
Ending: -s vs. -n	0.40	0.12	3.34	0.001

TABLE 7 Model summary for Percent Trochaic plural forms in the plural elicitation task with real words in L2 speakers (Task 2).

	Estimate	Std. error	z-value	p-value
(Intercept)	7.07	0.66	10.74	<0.001
Level1: Y1 vs. Y2	0.77	0.30	2.52	0.012
Level2: Y1+2 vs. Y3	1.73	0.32	5.47	<0.001
Level3: Y1-3 vs. Y3+	1.51	0.20	7.46	<0.001
Class: -el vs. -e	−5.25	0.74	−7.12	<0.001
Class: -en vs. -e	−3.48	0.78	−4.44	<0.001
Class: -er vs. -e	−4.83	0.73	−6.64	<0.001
Class: FinStress vs. -e	−1.73	0.72	−2.41	0.016
Class: Mono vs. -e	−1.21	0.75	−1.60	0.109
Class: StrChg vs. -e	−1.38	0.90	−1.53	0.127
Class: -V vs. -e	−2.28	0.75	−3.04	0.002
Ending: -e vs. -n	−1.72	0.17	−9.99	<0.001
Ending: -en vs. -n	−1.73	0.15	−11.31	<0.001
Ending: -er vs. -n	−1.94	0.34	−5.66	<0.001
Ending: -Ø vs. -n	−1.44	0.15	−9.37	<0.001
Ending: other vs. -n	−2.25	0.29	−7.85	<0.001
Ending: -s vs. -n	−2.33	0.18	−13.31	<0.001

in 0.57 (SD 0.50), 0.66 (SD 0.47), 0.75 (SD 0.43), and 0.82 (SD 0.39) of cases, respectively (cf. Supplementary Figure 5).

Performing the same binomial generalized linear mixed-effects model analysis as above with "Percent Trochaic" (trochaic vs. non-trochaic) as the binary dependent variable, a model comparison revealed that model fit improved when including (word) Class as a covariate ($\chi^2_{(7)} = 72.2$, $p < 0.001$), and further when including EndingChosen ($\chi^2_{(6)} = 317.3$, $p < 0.001$). The model output (see Supplementary Table 1) indicates that (proficiency) Level, Class, and EndingChosen were significant

FIGURE 6
Percentage of words produced with a particular ending (EndingChosen) in the plural elicitation task with real words (Task 2) in L2 speakers, split by (proficiency) Level and (word) Class.

predictors, as several of each of the Level, Class, and EndingChosen contrasts were significant. *Post-hoc* pairwise comparisons for Level showed that the percentage of trochaic plural forms was always significantly higher for higher Levels ($p < 0.02$), with two exceptions: The contrast between Year 1 and Year 2, and the contrast for Year 3 with and without immersion were not significant.

Overall, the results of just the incorrect plural forms suggest that the percentage of trochaic forms in the production of plural forms for existing German words increased with course level, although the two Levels at the low and high end of the proficiency spectrum (Year 1–Year 2, and Year 3–Year 3 + Imm., respectively) did not show statistically significant differences.

4.3 Results Task 3: grammatical acceptability judgment task (GAJT)

4.3.1 Correlations of ratings and plural categories

This task was designed as a well-formedness rating task of existing and made-up plural forms of existing German

words. First, we expected that incorrect forms would be rated as less acceptable than correct forms. The crucial question was whether the rating would reflect both the forms' accuracy and prosodic form such that within both the group of correct and the group of incorrect plurals, non-trochaic forms would be rated as less acceptable than trochaic forms. Thus, we hypothesized that sensitivities to both accuracy and the prosodic pattern in German plurals should be reflected in the ratings (as lower numbers on the scale indicate more acceptability) across the four tested accuracy-by-prosody Plural Types, in this order: Correct Trochaic < Correct Non-Trochaic < Incorrect Trochaic < Incorrect Non-Trochaic. Figure 7 presents the well-formedness rating results split by (proficiency) Level.

In the raw L1 and Year 3 (without immersion) data, it is apparent that incorrect plurals were less acceptable than correct plurals, and most importantly, that within incorrect plurals, the non-trochaic forms were less acceptable than the trochaic forms. It should be noted that the second group, the correct non-trochaic plurals, made up a very small category with just one type of plural, namely monosyllabic singular forms pluralized
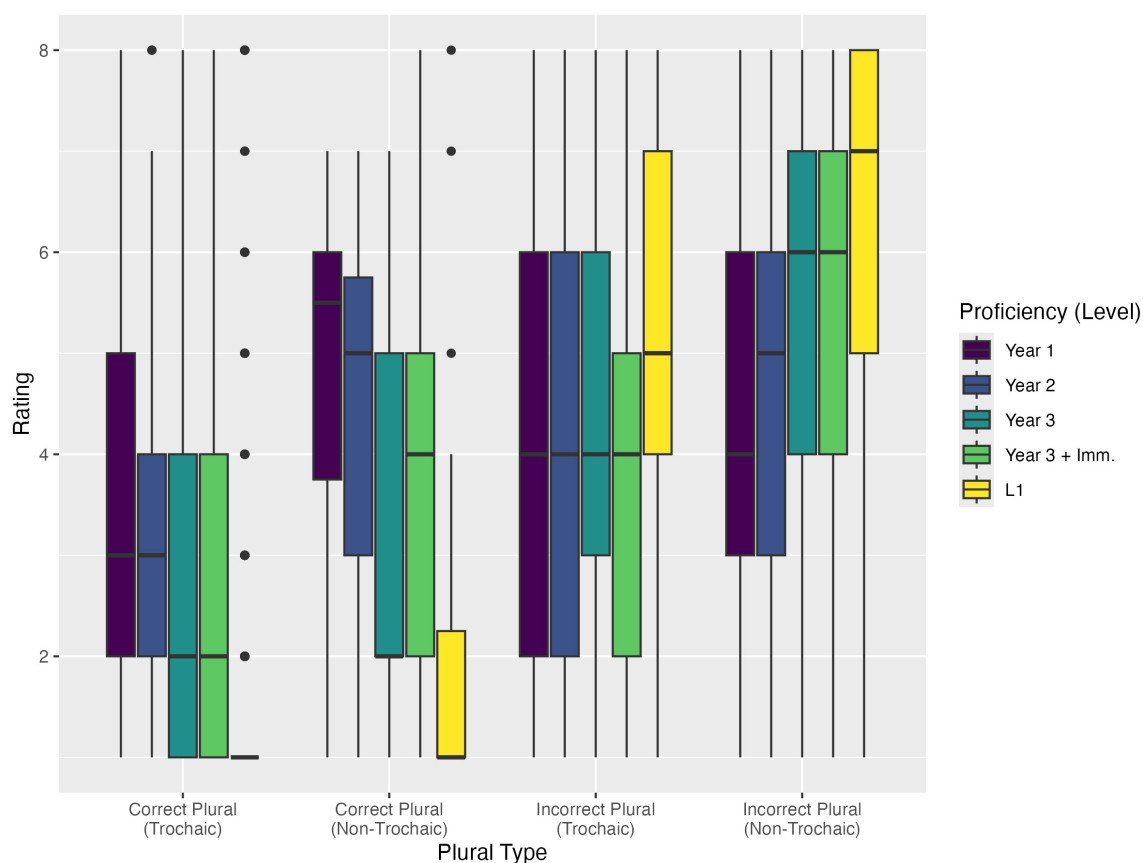
FIGURE 7
L2 and L1 users' well-formedness ratings (raw data) on the grammatical acceptability judgment task (GAJT, Task 3), split by Plural Type and (proficiency) Level.

with the -s suffix ($n = 6$).[18] Figure 7 presents, for each level of proficiency, a relatively steady increase in rating responses across all four Plural Types: Correct Plurals (Trochaic) < Correct Plurals (Non-Trochaic) < Incorrect Plurals (Trochaic) < Incorrect Plurals (Non-Trochaic), with the following exceptions for the second category of correct, non-trochaic plurals: The two mean values of Years 1 and 2 for the second category went against this overall trend, as they rejected non-trochaic but correct plurals strongly. Similarly, Year 3 + Imm. showed the same mean rating for the second (correct, non-trochaic plurals) and third category (incorrect, trochaic plurals). For the second Plural Type (correct, non-trochaic items), Year 1 and Year 2 users rejected these monosyllabic forms pluralized with the -s suffix on average more strongly than both types of incorrect plurals (Year 1), and more strongly than incorrect trochaic but approximately equally to incorrect non-trochaic plurals (Year 2).

We used the nonparametric Spearman's rank-order correlation test to analyze the ordinal Likert scale-like data. The four Plural Types were ranked in this order: Correct Trochaic—Correct Non-Trochaic—Incorrect Trochaic—Incorrect Non-Trochaic. We

tested whether the participants' ratings correlated with the four ranked types of plurals in the study at each of the five (proficiency) Levels (Year 1 through Year 3 + Imm., L1). The individual participants' ratings were first turned into by-participant ranking of scores (ascending order).[19] This type of data analysis helped address differences in individual participants' use of the rating scale, as visual inspection suggested that L2 users tended to use more of the central parts of the scale. On the other hand, L1 users as a group made use of the entire scale for the categories examined.

The results showed a moderate positive correlation for Year 1 and Year 2. For Year 1: Spearman's $r_s(1996)$: 0.343, $p < 0.001$; for Year 2: $r_s(2500)$: 0.451, $p < 0.001$. The Year 3 groups without and with immersion and the L1 users showed a strong positive correlation. For Year 3: $r_s(1662)$: 0.606, $p < 0.001$; for Year 3 + Imm.: $r_s(9999)$: 0.604, $p < 0.001$; for L1: $r_s(5038)$: 0.647, $p < 0.001$.

Thus, the analysis showed decreasing acceptability ratings at each (proficiency) Level for these four ranked Plural Types: Correct

_____

18   This class was included in the analysis at the request of a reviewer.

_____

19   This means that we computed the rank of each rating value within each participant's ratings; the numerically lowest rating was assigned rank 1, the next lowest rating was assigned rank 2, etc.

Plurals (Trochaic) < Correct Plurals (Non-Trochaic) < Incorrect Plurals (Trochaic) < Incorrect Plurals (Non-Trochaic). However, the first two Levels of L2 users, Years 1–2, showed a moderate correlation, while the higher (proficiency) Levels of the Year 3 users (with and without immersion) and the L1 users showed strong correlations.

4.3.2 Ratings by (word) Class and Ending

As for the other tasks, we next present the data broken down by Ending (suffix provided) in Figure 8 (the data are broken down by (word) Class in Supplementary Figure 6), to visually examine whether the ratings varied by this factor.

All three figures for the GAJT (Figures 7, 8, Supplementary Figure 6) illustrate that the L2 German users in this study assigned lower acceptability ratings to incorrect plural forms, including non-trochaic plural forms, with increasing proficiency. At the same time, more proficient users also rated correct items as more acceptable. This involved, in particular, learning to accept correct non-trochaic forms (the second Plural Type), i.e., correct monosyllabic plurals in -s. This group of correct plural forms was strongly rejected in Years 1 and 2; even Year 3 + Imm. rejected this type of correct plurals more strongly than some incorrect trochaic plurals. Additionally, L2 acquisition involves learning to accept correct trochaic plural forms which are marked for plurality with a null-suffix. Non-overtly marked plurals had the lowest acceptance rate among all correct trochaic forms in Year 1. In fact, across all L2 (proficiency) Levels, trochaic plurals with the null-suffix consistently received the lowest or one of the lowest ratings, both within the group of correct and within the group of incorrect plurals.

Finally, among incorrect forms, Year 1 participants showed the highest acceptability ratings for both trochaic and non-trochaic plural forms with the -en suffix. Year 1 participants' ratings reflected a preference for incorrect trochaic plurals with the -en suffix over correct trochaic plurals with a null-suffix. At Year 1, trochaic forms with the suffixes -en and -e showed higher mean ratings than their non-trochaic counterparts; meanwhile, both trochaic and non-trochaic plural forms with the -n suffix and the null-suffix (as well as the -s suffix) showed comparable mean acceptability ratings. For L2 users beyond Year 1, each individual suffix received lower ratings in non-trochaic compared to trochaic plurals.

5 Discussion

5.1 Answering the research questions

5.1.1 RQ1: prosodic structure of non-words (Task 1)

The first research question examines the extent to which L1 and L2 German speakers produced trochaic plural forms in the non-word plural elicitation task, i.e., Task 1. As expected, the L1 German speakers produced a large majority of trochaic plurals (0.92, SD 0.27). Perhaps more surprising is the large majority of trochaic plurals produced by the L2 users, ranging from 0.78 (SD 0.41) by the Year 1 group to 0.95 (SD 0.23) by the Year 3 group with immersion. Overall, there was indeed a significant effect

of (proficiency) Level, such that the more proficient the users, the more they produced trochaic plural forms. However, there was no significant difference between the Year 1 and Year 2 groups or between the two Year 3 groups (with and without immersion). Notably, both Year 3 groups actually produced percentages of trochaic plurals that were slightly higher—but not significantly different—than their L1 counterparts. This could point toward a slight tendency to overgeneralize the trochaic pattern in more advanced L2 users.

5.1.2 RQ2: accuracy and prosodic structure of real words (Task 2)

Recall that this set of research questions applies strictly to the L2 users of German and relates to their performance on a plural elicitation task using existing German words in terms of both accuracy (RQ2a) and creating trochaic plurals (RQ2b,c). In terms of accuracy, L2 users of German were not particularly successful at producing large numbers of accurate plurals overall, ranging from just 0.41 (SD 0.49) accuracy for participants in the Year 1 group to a high of 0.66 (SD 0.47) and 0.66 (SD 0.47) for users in the Year 3 groups, with and without immersion, respectively. Accuracy did improve significantly across proficiency Levels, although the two Year 3 groups did not differ significantly from each other.

In terms of the percentage of trochaic plurals produced during this task with real words overall (RQ1b) and on just incorrect words (RQ1c), a similar tendency was found. For all plurals produced as well as for just incorrect plural forms, the L2 users tended to produce more trochaic plural forms as proficiency increased (though without a significant difference between the Year 1 and Year 2 groups, and between the two Year 3 groups in each case). It is also worth noting that the mean percentage of trochaic forms was substantially higher than the means for each group's accuracy scores, although it has to remain acknowledged that the actual values ranged both above and below that mean (e.g., Year 1 means: trochaic 0.73 vs. accuracy 0.41—difference: 31.8 percentage points; Year 3 + Imm. means: trochaic 0.92 vs. accuracy 0.66—difference: 25.6 percentage points). This indicates—and was confirmed when examining just forms that were incorrectly pluralized—that even when L2 users formed incorrect plurals, those forms became more trochaic as proficiency increased.

5.1.3 RQ3: prosodic structure and well-formedness judgments (Task 3)

Results from the well-formedness judgment task (Task 3) revealed that there were significant correlations between the participants' ratings and the four types of plural forms which differed in accuracy (correct/incorrect) crossed with prosodic structure (trochaic/non-trochaic). This indicates that, overall, L1 and L2 German users rated non-trochaic forms as less well-formed compared to trochaic forms, both within the group of correct and within the group of incorrect plural forms provided. Furthermore, the strength of the correlations differed between (proficiency) Levels. L1 users and L2 users in Year 3 (with and without immersion) showed a strong correlation, while less proficient L2 users in Years 1 and 2 showed only a moderate correlation.
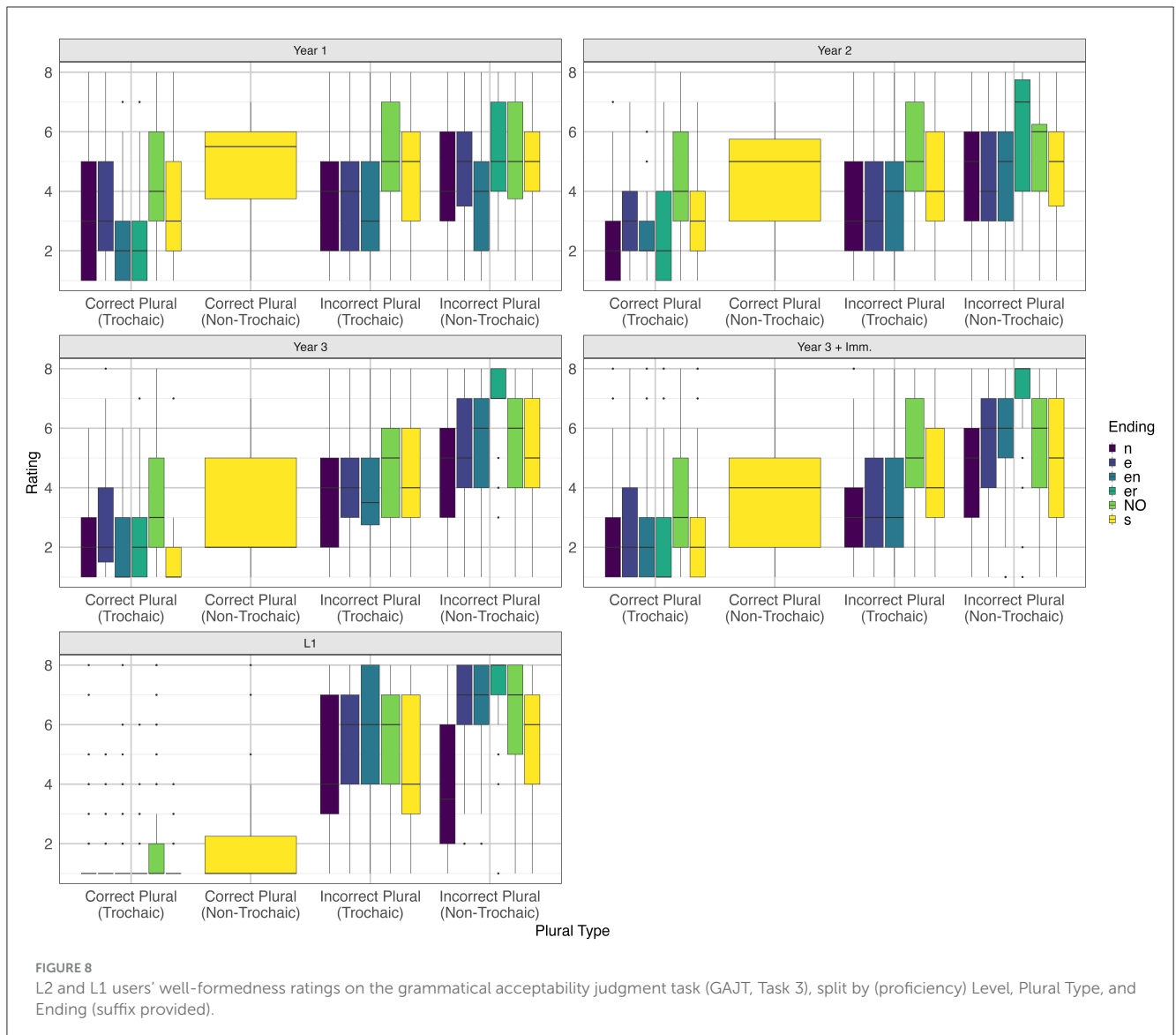
FIGURE 8
L2 and L1 users' well-formedness ratings on the grammatical acceptability judgment task (GAJT, Task 3), split by (proficiency) Level, Plural Type, and Ending (suffix provided).

5.1.4 Summary of the experimental data

One key take-away from the three tasks in this study is that L1 English-L2 German users produced more trochaic plurals and rated plural forms based on both accuracy and prosodic structure with increasing experience and proficiency. The data presented above also suggest that this group of L2 users started out with high levels of trochaic plural forms, around 0.73–0.78 for real and non-words, respectively. More advanced L2 users may have even tended to overgeneralize the trochaic pattern in the non-word data (although these differences were not statistically significant). The L2 speakers' performance in Task 2 with plural elicitations for existing German words indicated that L2 users might have been relatively more adept at the prosodic pattern compared to their overall accuracy for plural markings. (It should be acknowledged, however, that the potential for non-target-like performance on plural accuracy is naturally several times higher than for non-target-like performance on prosodic shape.)

5.2 Implications for formal models and accounts of German plural (acquisition)

5.2.1 The role of prosody in L1 and theoretical models

The results regarding the prosodic pattern from the three tasks outlined above are in line with previous experimental studies (e.g., Domahs et al., 2013, 2017; Kauschke et al., 2013) reviewed in the background section. As a reminder, Kauschke et al. (2013) noted that German-speaking children without language impairments produced more trochaic plural forms for real and non-words than their peers with impairments. Further clinical evidence in favor of the "optimal prosodic shape" in perception, i.e., in the interpretation of forms as singular or plural, was provided by a case study on a patient with aphasia—a person with impaired lexical knowledge (Domahs et al., 2017).

A number of theoretical analyses of the German plural, including those reviewed in Section 2, have drawn on the trochee

as an important factor or constraint in shaping German plurals. The empirical results in this study are in line with theoretical accounts which analyze German plurals as containing a word-final disyllabic trochee. In particular, they are in line with generative accounts in which the trochee is modeled as a constraint (Wiese, 2009) or template (Smith, 2004, 2020; cf. Schuhmann and Putnam, 2021 for a phase-based account within Distributed Morphology) in the grammar for German plurals. From an OT perspective, in particular, it could be argued that during their language development, L2 learners of German come to rank a well-formedness constraint for trochees undominated in its area of application or weighted more heavily than other constraints related to the morphophonology of German plural formation (cf. Pater, 2009). Yet, detailed analyses of how the prosodic pattern interacts with morphology allow us to also evaluate other central models discussed in the background section.

5.2.2 The interplay of prosody and morphology

An exploratory look at the interplay of the prosodic patterns with the morphology, i.e., the stimuli's (word) Class and EndingChosen (in the two elicitation tasks) or Ending (suffix provided in the rating task) revealed several important aspects. First, the suffix -s was rarely used in the elicitation tasks (cf. also Schuhmann and Smith, accepted); when it was used, it mostly occurred on words ending in a syllable with an unstressed full vowel, both in L1 and L2 speakers across (proficiency) Levels. This finding in particular would not have been predicted by (the original version of) the dual route models (Clahsen, 1999; Marcus et al., 1995). As described in Section 2, in the dual route models, the -s suffix is considered by some scholars to be the default and would have been expected to occur very frequently with new or unknown words.

Secondly, the morphological analyses revealed which specific contexts correlated with lower percentages of trochaic forms in the elicitation tasks or lower ratings in the GAJT among the less proficient L2 users. Here, we found that L2 users started out with a very low percentage of trochaic plural forms for singulars ending in closed schwa-syllables. This resulted from learners in Year 1 frequently using syllabic suffixes instead of null-suffixes to mark plurality on these words (except for words ending in -en), most notably the suffix -en, less frequently -e, and occasionally -er. Null-marking gradually increased over the subsequent (proficiency) Levels but still did not reach L1 levels by Year 3. In place of the null-suffixes for words in closed schwa-syllables, L2 users produced more -n suffixes through Year 3 + Imm.; this led to forms which maintain a trochee but are segmentally not target-like. As a result, the percentage of trochaic plural forms increased across (proficiency) Levels for words in closed schwa-syllables, while segmentally, the plural forms still differed markedly from the L1 group by Year 3.

5.2.3 A proposal of relevant factors in the acquisitional path of German plurals

We suggest that the pattern discussed above for words ending in closed schwa-syllables among early L2 German learners can be accounted for by a few driving forces for

marking the function of plurality on nouns in early L2 acquisition. These forces have already been proposed as part of Köpcke's schema model discussed above (e.g., Köpcke, 1988, inter alia). As a brief reminder, Köpcke's schema model proposes that certain plural forms signify plurality more strongly than other plural forms, depending on the strength of the cues, the latter of which include saliency, frequency, validity, and iconicity.

First, the patterns in the data overall can be captured with the notion of iconicity, i.e., a principle which marks plurality explicitly and overtly, thus rendering plurals different from singulars (cf. Eisenberg, 2020; Köpcke, 1988 for slightly different definitions). Secondly, Year 1 speakers frequently add syllabic suffixes such as -en, which could additionally be captured by the related strategy of saliency in Köpcke's schema model (e.g., Köpcke, 1988, inter alia). This expresses the notion that syllabic suffixes serve as a better cue to plurality than non-syllabic suffixes or umlaut. In the GAJT, the L2 users in Year 1 rated plurals formed with the suffix -en better than those formed with a null-suffix in all relevant Plural Types. This further corroborates the findings from the elicitation tasks where L2 learners initially overgeneralized the -en suffix to contexts where a null-suffix was expected.

Thirdly, in the GAJT (Task 3), learners in their first years of study strongly rejected existing monosyllabic plurals formed with the suffix -s. Together with the trisyllabic plurals for words ending in closed schwa-syllables just discussed, the L2 learners seemed to start out by requiring multisyllabic plural forms—in addition to iconicity and saliency. All three of these principles have been proposed in Köpcke's schema model (e.g., Köpcke, 1988, inter alia). These notions overlap partially, meaning that it is not always possible to identify which principle is the driving force in the observable behavior. Note, however, that this is very much in the spirit of Köpcke's schema model, in which prototypical plural forms unite a maximal number of cues to plurality (such as iconicity, saliency, multisyllabicity or—in later versions—trochee) on themselves.

It is also worthwhile to consider possible transfer effects from the learners' L1. As it stands, the contribution of iconicity might be reinforced by the learners' L1 patterns, as English plurals are—with a few exceptions—typically distinct from their singular forms. On the other hand, multisyllabicity is not a prevalent pattern in L1 English plural formation and may thus be less likely to be reinforced by L1 patterns, and therefore appears to be a strategy that the learners developed specifically for German plurals. Overall, the data from the early L2 learners can be meaningfully analyzed using the strategies proposed in the schema model by Köpcke and colleagues.

5.2.4 Comparison with other data on L2 German plural acquisition

The fact that German learners in Year 1 produced a substantial percentage of non-trochaic forms—albeit only for the closed schwa-syllables—is intriguing since other work on L2 German acquisition reported only a handful of cases of prosodically ill-formed plurals. Previous studies on L2 acquisition that have examined prosody were either concerned with children learning German as a second

language in schools in Germany (Köpcke and Wecker, 2017; Wecker, 2016) or with adult foreign language learners of German at an Italian university (pilot study by Vogt, 2016). Children with SLI, however, were reported to form prosodically non-optimal plurals (Kauschke et al., 2011, but see Kauschke et al., 2013). While our study includes data from adult foreign language learners of German in their very early L2 development without much target-language input, both the child and adult L2 learners in the literature had received many years of input or exposure to German. This suggests that with increased target-language input, non-clinical learners will come to acquire the specific prosodic pattern of German plurals, i.e., trochaic rather than multisyllabic plural forms, as is also the case in the study reported here.

In order to account for the developmental path across proficiency levels in child L2 acquisition, an adapted schema model with a second-order schema has recently been proposed (Köpcke and Wecker, 2017; Wecker, 2016). These studies argue that second-order schemata, which are intended to capture paradigmatic relationships such as the contrast between singular and plural forms, are acquired late in child L2 German learners. For our data with adult foreign language learners, on the other hand, the paradigmatic contrast—which, as indicated above, overlaps with the drive for iconicity, saliency, and multisyllabicity—appears strongest at the very beginning of the language acquisition process. Future research should continue to investigate how the acquisitional path in the data presented here could be accounted for with a further revised schema model.[20]

5.2.5 The relevance of prosody in acquisition and theoretical models

Overall, the interplay of morphology and prosody in our exploratory analysis suggests that prosody plays a central role in the acquisitional path in our data. The role of prosody materializes in several forms: There appears to be an initial requirement for *multisyllabic* plurals (thus rejecting monosyllabic plurals in -*s*), as well as saliency, here defined as adding a *syllabic* suffix to a singular form to mark plurality. Thus, the nuanced analyses of the (word) Classes show that prosody plays a foundational role in the plural formation from the beginning, even if it might not lead to target-like trochaic productions or ratings, and even if it might not be the only factor involved in plural formation.

Yet, in the end, the available data to date do not allow us to unambiguously decide between various theoretical models and accounts of German plural formation or acquisition. Crucially, the data from Year 1 and subsequent (proficiency) Levels suggest that syllabic suffixes—notably -*en*—on words ending in closed schwa-syllables reduce with increasing proficiency, making way for more null-suffixes and, in particular, -*n* suffixes. On the one hand, this could be accounted for by a trochaic pattern for plurals becoming stronger with increasing proficiency. On the other hand, this could also be accounted for by L2 users learning the distribution of

suffixes, i.e., which types of (word) Classes co-occur with which plural suffixes across (proficiency) Levels. This latter account would not necessarily require a prosodic condition to capture the same pattern of data. The trochaic pattern in L1 and especially more advanced L2 users may then be a mere by-product (e.g., Trommer, 2021) or just one of many cues to plurality (see Köpcke and colleagues' schema model, e.g., Köpcke, 1988; Köpcke and Wecker, 2017). We have sketched an initial proposal of how the acquisitional path could be captured with the strategies for plural marking in the schema model, although it appears that the relevance of specific cues and their interaction would have to be different than the proposed model for child L2 acquisition (Köpcke and Wecker, 2017; Wecker, 2016).

5.2.6 Limitations and open questions

Readers should keep in mind that the data presented in this study are cross-sectional. Ideally, further research would add a longitudinal perspective to the cross-sectional data presented here to follow the same learners as they develop their L2 language skills over time. This kind of work could then also test whether and how these prosodic patterns develop within individuals over the course of their L2 acquisition process and how this might align with the PTH (cf. also Cabrelli, 2019). Such a study could provide insights into individual acquisitional paths and individual differences related to both the production and perceptivity of prosodic patterns in plural forms, and their potential interaction with other plural markings, specifically, umlaut and suffixal choice for plurals. Additional caution needs to be taken when interpreting the results reported above due to imbalances in the data. The analyses of the interaction of prosody and morphology—(word) Class and Ending/EndingChosen—presented here should be considered exploratory and will need to be ratified in future work.

Finally, while not explicitly tested here, the similarity between participants' L1 English and the target language German may contribute to facilitating effects that may not necessarily be replicated with L2 users whose L1 is prosodically different from German. For instance, English-speaking and French-speaking adult L2 German learners differ in their preferences for lexical stress assignment based on their L1 (O'Brien and Sundberg, 2023). More cross-linguistic work in this area could be another testing ground for the role of prosody and perhaps further examine the validity of accounts with prosodic constraints or templates. English and German share how prosodic prominence is used to mark word stress, and the trochee is a prevalent pattern in German, English, and Dutch (cf. Domahs et al., 2014). Other L1 language backgrounds might include pitch-accent languages, tonal languages, or a language that might not utilize (trochaic) feet, as has been suggested for Portuguese (cf. Garcia and Goad, 2021). Garcia and Goad (2021) argue that, while metrical stress data for English align with a foot-based analysis, Portuguese metrical stress is not captured "optimally" with an analysis that assumes feet. The authors provide additional evidence from sonority effects and word minimality issues from both languages in support of their analysis that some languages may not build feet. Thus, future research could investigate L1 Portuguese learners of German to determine whether this group of L2 users would

---

20 In fact, words ending in closed schwa-syllables are central to the schema model and present one promising avenue for further research among our adult L2 users.

show patterns that align with a foot-based prosodic analysis of German plurals.

6 Conclusions

Although many descriptive and theoretical accounts of German plurals have drawn on the syllabic or prosodic structures of nouns, there have hitherto only been a few empirical studies testing the productivity of this prosodic aspect or the interaction of prosody and morphology for German plurals in clinical or acquisitional contexts (but see Domahs et al., 2017; Kauschke et al., 2011, 2013; Vogt, 2016). The data presented in this study arguably confirm that the word-final trochee requirement is a productive pattern in L1 users of German. Similarly, the trochaic pattern progressively developed as a productive pattern in the plural elicitation data from L1 English-L2 German speakers: As proficiency increased, L2 users produced more trochaic plural forms, both in non-words and in existing words of the German lexicon, whether pluralized accurately or not. The descriptive data from the rating study further suggest that this prosodic structure in plurals develops concomitantly as learners progress in their L2 proficiency. In fact, throughout the study tasks, by the third year of university study, L2 users produced and preferred (in their ratings) trochaic plural forms.

To this end, the results of our behavioral psycholinguistic study are consistent with an account in which the trochaic template for German nominal plural formation is part of L1 users' grammars and mental representations, and develops with increasing proficiency in L1 English-L2 German participants' grammars. Yet, based on the available data, we cannot unambiguously rule out accounts of German plurals in which the prosodic pattern is merely epiphenomenal or a by-product of morphological patterns (e.g., Trommer, 2021) rather than a prosodic constraint or principle that learners need to acquire separately from the morphology. In the latter case, learners might still produce and prefer trochaic forms—or develop these with increasing input and target language proficiency—but without the need for a separate constraint or template. Our exploratory analyses of the interplay between prosody and morphology of German plurals suggest that iconicity, saliency, and multisyllabicity— factors from Köpcke's schema model (i.e., Köpcke, 1988, *inter alia*)—could explain the early phases of the acquisitional path in the adult foreign language learning data presented here.

We leave it open for future psycholinguistic research, and perhaps computational modeling, to further examine whether L1 and L2 language users develop sensitivities for the prosodic patterning itself, which would be in line with generative accounts (e.g., Schuhmann and Putnam, 2021; Smith, 2004, 2020; Wiese, 1996, 2009), or whether users primarily develop sensitivities for the distribution of the plural allomorphs. In the end, we hope that the findings presented here invite further cross-linguistic inquiries into the development of prosodic patterns in the acquisition of grammar in L1 and L2 users and into how prosodic cues and morphology or other higher-level linguistic structures interact during L2 development in various language learning contexts.

Data availability statement

The datasets analyzed for this study can be found in the Open Science Framework (osf.io): https://osf.io/nfpb5/. Further inquiries can be directed to the corresponding author KS.

Ethics statement

The studies involving humans were approved by Institutional Review Board (IRB) at Brigham Young University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of

their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/flang.2024.1338625/full#supplementary-material

References

Archibald, J. (2022). "Feature dependency and the poverty of the stimulus in the acquisition of L2 German plural allomorphy," in *Generative SLA in the Age of Minimalism: Features, interfaces, and beyond. Selected Proceedings of the 15th Generative Approaches to Second Language Acquisition Conference*, eds. T. Leal, E. Shimanskaya, and C. A. Isabelli (Amsterdam: John Benjamins), 117–136.

Augst, G. (1979). Neuere Forschungen zur Substantivflexion. *Zeitschrift für Germanistische Linguistik* 7, 220–232.

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Memory Lang.* 59, 390–412. doi: 10.1016/j.jml.2007.12.005

Baayen, R. H., and Schreuder, R. (1999). War and peace: morphemes and full forms in a non-interactive activation parallel route model. *Brain Lang.* 68, 27–32. doi: 10.1006/brln.1999.2069

Bates, D. M, Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear mixed-effects models using lme4. *J. Statist. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Beretta, A., Campbell, C., Carr, T. H., Huang, J., Schmitt, L. M., Christianson, K., et al. (2003). An ER-fMRI investigation of morphological inflection in German reveals that the brain makes a distinction between regular and irregular forms. *Brain Lang.* 85, 67–92. doi: 10.1016/S0093-934X(02)00560-6

Bermúdez-Otero, R. (2018). "Stratal phonology," in *The Routledge Handbook of Phonological Theory*, eds. S. J. Hannahs and A. R. K. Bosch (Abingdon: Routledge), 100–134.

Binanzer, A., and Wecker, V. (2020). "Lernergrammatik und Grammatikunterricht – Überlegungen zu natürlichen Erwerbssequenzen und didaktischer Progression," in eds. L. Miriam *Grammatikdidaktik – theoretische und empirische Zugänge zu sprachlicher Heterogenität* (Baltmannsweiler: Schneider), 61–85.

Bittner, D. (1991). *Von starken Feminina und schwachen Maskulina: die neuhochdeutsche Substantivflexion; eine Systemanalyse im Rahmen der natürlichen Morphologie* (PhD diss.). Universität Jena, Jena, Thuringia, Germany.

Booij, G. (1998). "Phonological output constraints in morphology," in *Phonology and Morphology of the Germanic Languages*, eds. W. Kehrein, and R. Wiese (Tübingen: Max Niemeyer Verlag), 143–163.

Booij, G. (2002). "Prosodic restrictions on affixation in Dutch," in *Yearbook of Morphology 2001*, eds. G. Booij and J. van Marle (Dordrecht: Springer), 183–201.

Cabrelli, J. (2019). Plotting individual learning trajectories in the acquisition of L2 prosodic constraints. *Linguist. Approaches Biling.* 9, 822–826. doi: 10.1075/lab.19065.cab

Clahsen, H. (1999). Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behav. Brain Sci.* 22, 991–1013. doi: 10.1017/S0140525X99002228

Clahsen, H., Eisenbeiss, S., and Sonnenstuhl-Henning, I. (1997). Morphological structure and the processing of inflected words. *Theoret. Linguist.* 23, 201–250. doi: 10.1515/thli.1997.23.3.201

Clahsen, H., Rothweiler, M., Woest, A., and Marcus, G. F. (1992). Regular and irregular inflection in the acquisition of German noun plurals. *Cognition* 45, 225–255. doi: 10.1016/0010-0277(92)90018-D

Clahsen, H., Sonnenstuhl, I., and Blevins, J. (2003). "Derivational morphology in the German mental lexicon: A dual mechanism account," in *Morphological Structure in Language Processing*, eds. R. Baayen and R. Schreuder (Berlin, New York: De Gruyter Mouton), 125–156.

Demuth, K. (1994). "On the 'underspecification' of functional categories in early grammars," in *Syntactic Theory and First Language Acquisition: Cross-Linguistic Perspectives*, eds. B. Lust, M. Suñer, and J. Whitman (Hillsdale, NJ: Lawrence Erlbaum Associates), 119–34.

Demuth, K. (2009). "The prosody of syllables, words and morphemes," in *Cambridge Handbook on Child Language,* ed. E. Bavin (Cambridge: Cambridge University Press), 183–198.

Demuth, K., and McCullough, E. (2009). The prosodic (re)organization of children's early English articles, *J. Child Lang.* 36, 173–200. doi: 10.1017/S0305000908008921

Domahs, F., Bartha-Doering, L., Domahs, U., and Delazer, M. (2017). "Wie muss ein guter deutscher Plural klingen?," in *Sichtbare und hörbare Morphologie,* eds. N. Fuhrhop, R. Szczepaniak, and K. Schmidt (Berlin, Boston: De Gruyter), 205–236.

Domahs, U., Lohmann, K., Moritz, N., and Kauschke, C. (2013). The acquisition of prosodic constraints on derivational morphology in typically developing children and children with SLI. *Clini. Linguist. Phonet.* 27, 555–573. doi: 10.3109/02699206.2013.796405

Domahs, U., Plag, I., and Carroll, R. (2014). Word stress assignment in German, English and Dutch: quantity-sensitivity and extrametricality revisited. *J. Comparat. Germa. Linguist.* 17, 59–96. doi: 10.1007/s10828-014-9063-9

Duden-Deutsches Universalwörterbuch (2023). Dudenredaktion (Ed.). *Duden. Deutsches Universalwörterbuch. 10., völlig überarbeitete und erweiterte Auflage.* Berlin: Dudenverlag.

Duden-Grammatik (1995). Drosdowski, Günther (Ed.). *Duden. Grammatik der deutschen Gegenwartssprache. 5., völlig neu bearbeitete und erweiterte Auflage.* Mannheim: Dudenverlag.

Duden-Grammatik (1998). Dudenredaktion (Ed.). *Duden. Grammatik der deutschen Gegenwartssprache. 6., neu bearbeitete Auflage.* Mannheim: Dudenverlag.

Eisenberg, P. (2020). *Grundriss der deutschen Grammatik. Band 1: Das Wort. 5. Aufl.* Stuttgart: Metzler.

Féry, C. (1994). Umlaut and inflection in German. *ROA-34, Rutgers Optimality Archive.* Available at: http://roa.rutgers.edu/files/34-1094/34-1094-F%E9RY-0-0.PDF (accessed August 31, 2024).

Garcia, G. D., and Goad, H. (2021). Weight Effects and the Parametrization of the Foot: English vs. Portuguese. Available at: https://osf.io/pc6hg/download (accessed August 31, 2024).

Gerken, L. (1996). Prosodic Structure in Young Children's Language Production. *Language* 72, 683–712. doi: 10.2307/416099

Gleason, J. B. (1958). *The Wug Test.* Cambridge: Larchwood Press.

Goad, H., and White, L. (2019). Prosodic effects on L2 grammars. *Linguist. Approaches Biling.* 9, 769–808. doi: 10.1075/lab.19043.goa

Hahn, U., and Nakisa, R. C. (2000). German inflection: single route or dual route? *Cognit. Psychol.* 41, 313–360. doi: 10.1006/cogp.2000.0737

Hall, T. A. (1992). *Syllable structure and syllable related processes in German.* Tübingen: Niemeyer.

Itô, J., and Mester, A. (1997). "Sympathy Theory and German truncations," in *Proceedings of the Hopkins Optimality Workshop/Maryland Mayfest 1997*, eds. by V. Miglio and B. Moren (College Park, MD: University of Maryland), 117–138.

Kauschke, C., Kurth, A., and Domahs, U. (2011). Acquisition of German noun plurals in typically developing children and children with specific language impairment. *Child Dev. Res.* 2011:718925.

Kauschke, C., Renner, L., and Domahs, U. (2013). Prosodic constraints on inflected words: an area of difficulty for German-speaking children with specific language impairment? *Clini. Linguist. Phonet.* 27, 574–593. doi: 10.3109/02699206.2013.798033

Kiparsky, P. (2015). "Stratal OT: A synopsis and FAQs," in *Capturing Phonological Shades*, eds. E. Y. Hsiao and L. Wee (Cambridge: Cambridge Scholars Publishing), 2–44.

Köpcke, K.-M. (1988). Schemas in German plural formation. *Lingua* 74, 303–335. doi: 10.1016/0024-3841(88)90064-2

Köpcke, K.-M. (1993). *Schemata bei der Pluralbildung im Deutschen.* Tübingen: Narr.

Köpcke, K.-M. (1998). The acquisition of plural marking in English and German revisited: schemata versus rules. *J. Child Lang.* 25, 293–319. doi: 10.1017/S0305000998003407

Köpcke, K. M., Schimke, S., and Wecker, V. (2021). Processing of German noun plurals: Evidence for first-and second-order schemata. *Word Struct.* 14, 1–24. doi: 10.3366/word.2021.0173

Köpcke, K. M., and Wecker, V. (2017). Source-and product-oriented strategies in L2 acquisition of plural marking in German. *Morphology* 27, 77–103. doi: 10.1007/s11525-016-9292-5

Laaha, S. (2011). Sonority, gender and the impact of suffix predictability on the acquisition of German noun plurals. *Lang. Interact. Acquisit.* 2, 82–100. doi: 10.1075/lia.2.1.04laa

Laaha, S., Ravid, D., Korecky-Kröll, K., Laaha, G., and Dressler, W. U. (2006). Early noun plurals in German: regularity, productivity or default? *J. Child Lang.* 33, 271–302. doi: 10.1017/S0305000906007379

Lenth, R. V. (2021). *emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.7.1-1.* Available at: https://CRAN.R-project.org/package=emmeans (accessed August 31, 2024).

Lleó, C., and Demuth, K. (1999). "Prosodic constraints on the emergence of grammatical morphemes: Crosslinguistic evidence from Germanic and Romance languages," in *Proceedings of the 23rd Annual Boston University Conference on Language Development, Vol. 2*, eds. A. Greenhill, H. Littlefield, and C. Tano (Cascadilla Press), 407–418.

Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., and Pinker, S. (1995). German inflection: the exception that proves the rule. *Cogn. Psychol.* 29, 189–256. doi: 10.1006/cogp.1995.1015

McCarthy, J., and Prince, A. (1994). "The emergence of the unmarked: Optimality in Prosodic Morphology," in *Proceedings of the North East Linguistics Society,* ed. M. Gonzàlez (Amherst, MA: Graduate Linguistic Student Association), 333–379. doi: 10.7282/T3Z03663

McCarthy, J. J., and Prince, A. (1996). "Prosodic morphology 1986," in *Linguistics Department Faculty Publication Series*, 13. Available at: http://ruccs.rutgers.edu/roa.html (accessed August 31, 2024).

Mugdan, J. (1977). *Flexionsmorphologie und Psycholinguistik. Untersuchungen zu sprachlichen Regeln und ihrer Beherrschung durch Aphatiker, Kinder und Ausländer, am Beispiel der deutschen Substantivdeklination.* Tübingen: Gunter Narr.

Neef, M. (1998). "The reduced syllable plural in German," in *Models of Inflection*, eds. R. Fabri, A. Ortmann, and T. Parodi (Tübingen: Niemeyer), 244–265.

Noske, R. (1993). *A theory of syllabification and segmental alternation: With studies on the phonology of French, German, Tonkawa, and Yawelmani.* Tübingen: Niemeyer.

O'Brien, M., and Sundberg, R. (2023). Lexical stress assignment preferences in L2 German. *Int. Rev. Appl. Linguist. Lang. Teach.* 61, 449–478. doi: 10.1515/iral-2020-0104

Pater, J. (2009). Weighted constraints in generative linguistics. *Cognit. Sci.* 33, 999–1035. doi: 10.1111/j.1551-6709.2009.01047.x

Penke, M., and Krause, M. (2002). German noun plurals: a challenge to the dual-mechanism model. *Brain Lang.* 81, 303–311. doi: 10.1006/brln.2001.2526

Pinker, S., and Ullman, M. T. (2002). The past and future of the past tense. *Trends Cogn. Sci.* 6, 456–463. doi: 10.1016/S1364-6613(02)01990-3

Prince, A. and Smolensky, P. (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar.* Hoboken, NJ: John Wiley.

R Core Team (2021). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/ (accessed August 31, 2024).

Revithiadou, A. (2007). "Colored turbid accents and containment: a case study from lexical stress," in *Freedom of Analysis?*, eds. S. Blaho, P. Bye, and M. Krämer (Berlin: De Gruyter), 149–174.

Schreuder, R., and Baayen, R. H. (1995). "Modeling morphological processing," in *Morphological Aspects of Language Processing,* L. B. Feldman (Hillsdale, NJ: Erlbaum), 131–154.

Schuhmann, K. (2015). "Degrees of (un)acceptability in syllable contact," in *Rutgers Optimality Archive*, 1–30. Available at: http://roa.rutgers.edu/article/view/1496 (accessed August 31, 2024).

Schuhmann, K. S., and Putnam, M. (2021). Relativized prosodic domains: a late-insertion account of German plurals. *Languages* 6, 1–20. doi: 10.3390/languages6030142

Schuhmann, K. S. and Smith, L. C. (accepted). "Interindividuelle Unterschiede in der Pluralallomorphie von Kunstwörtern in L1 und L2 Deutsch-Sprachverwendenden," in *Beiträge zur Fremdsprachenforschung (BFF) 18. (Dokumentation zum 30. Kongress für Fremdsprachendidaktik der Deutschen Gesellschaft für Fremdsprachenforschung (DGFF), Freiburg 2023)*, eds. N. Kulovics, O. Mentz, and T. Raith (Bielefeld: Schneider bei wbv).

Schuhmann, K. S., and Smith, L. C. (2022). Practical prosody: a new hope for teaching German plurals. *Die Unterrichtspraxis/Teaching German* 55, 1–24. doi: 10.1111/tger.12192

Smith, L. C. (2004). *Cross-Level Interactions in West Germanic Phonology and Morphology* (Dissertation). University of Wisconsin–Madison, Madison, WI, United States.

Smith, L. C. (2020). "The role of foot structure in Germanic," in *Cambridge Handbook of Germanic Linguistics*, eds. R. Page, and M. Putnam (Cambridge: Cambridge University Press), 49–72.

Smith, L. C. (2022). "Moving from syllables to feet in the prosodic hierarchy: how foot-based templates reflect prosodic preferences," in *Linguistic Preferences*, eds. P. Noel Aziz Hanna and L. Smith (Berlin, Boston: De Gruyter Mouton), 109–140.

Smith, L. C. (2024). A case for Old Frisian vowel balance as prosodically-driven vowel reduction. *Amsterdamer Beiträge zur älteren Germanistik.* 84, 372–394. doi: 10.1163/18756719-12340328

Sonnenstuhl, I., Eisenbeiss, S., and Clahsen, H. (1999). Morphological priming in the German mental lexicon. *Cognition* 72, 203–236. doi: 10.1016/S0010-0277(99)00033-5

Sonnenstuhl-Henning, I. (2010). *Deutsche Plurale im mentalen Lexikon: Experimentelle Untersuchungen zum Verhältnis von Speicherung und Dekomposition.* Tübingen: Max Niemeyer Verlag.

Szagun, G. (2001). Learning different regularities: The acquisition of noun plurals by German speaking children. *First Lang.* 21, 109–141. doi: 10.1177/014272370102106201

Trommer, J. (2021). The subsegmental structure of German plural allomorphy. *NLLT* 39, 601–656. doi: 10.1007/s11049-020-09479-7

van der Hulst, H., and Kooij, J. (1998). "Prosodic choices and the Dutch nominal plural," in *Phonology and Morphology of the Germanic Languages*, eds. W. Kehrein, and R. Wiese (Tübingen: Max Niemeyer Verlag), 187–197.

van Oostendorp, M. (2008). Incomplete devoicing in formal phonology. *Lingua* 118, 1362–1374. doi: 10.1016/j.lingua.2007.09.009

Vogt, B. (2016). Prosodiegeleitete Grammatik: Zum Erwerb der Pluralflexion im Deutschen. *German Foreign Language* 2, 67–80. Available at: http://gfl-journal.de/wp-content/uploads/2024/06/20160206-vogt.pdf (accessed August 31, 2024).

Wecker, V. (2016). "Strategien bei der Pluralbildung im DaZ-Erwerb," in *Eine Untersuchung mit russisch und türkischsprachigen Lernern*. Berlin: De Gruyter.

Wegener, H. (1995). *Die Nominalflexion des Deutschen, verstanden als Lerngegenstand*. Tübingen: Max Niemeyer Verlag.

Wiese, R. (1986). Schwa and the structure of words in German. *Linguistics* 24, 697–724.

Wiese, R. (1996). *The Phonology of German*. Oxford: Oxford University Press.

Wiese, R. (2009). The grammar and typology of plural noun inflection in varieties of German. *J. Comparat. German. Linguist.* 12, 137–173. doi: 10.1007/s10828-009-9030-z

Wunderlich, D. (1999). *German Noun Plural Reconsidered*. Düsseldorf: University of Düsseldorf. Available at: https://user.phil.hhu.de/~wdl/Nplural.pdf (accessed August 31, 2024).

Wurzel, W. (1984). *Flexionsmorphologie und Natürlichkeit*. Berlin: Akademie-Verlag.

Wurzel, W. U. (1998). "Drei Ebenen der Struktur von Flexionsparadigmen," in *Models of Inflection*, eds. R. Fabri, A. Ortmann, and T. Parodi (Tübingen: Niemeyer), 225–243.

Zaretsky, E., Neumann, K., Euler, H., and Lange, B. (2013). Pluralerwerb im Deutschen bei russisch- und türkischsprachigen Kindern im Vergleich mit anderen Migranten und monolingualen Muttersprachlern. *Zeitschrift für Slawistik* 58, 43–71. doi: 10.1524/slaw.2013.0003

Zimmermann, E. (2014). *A Phonological Account of Morphological Length* (PhD diss.) Universität Leipzig, Leipzig, Germany.

Zimmermann, E. (2017). *Morphological Length and Prosodically Defective Morphemes*. Oxford: Oxford University Press.