

Morphosemantic Role of the Regular Plural Morpheme Within English Compounds

Amira Abdullah Alshehri

Department of Languages and Translation, University of Tabuk, Tabuk, Saudi Arabia

Abstract—This study investigates the semantic role of the regular plural morpheme in the modifiers of English compounds (e.g., *publications catalogue*, *weapons inspection*), specifically examining whether this morpheme functions solely as a grammatical connector or also contributes semantic meaning. Sixty-seven monolingual native English speakers rated the acceptability of 80 sentences containing novel compounds with a medial regular plural morpheme. The compounds were divided into two types: those referring to many tokens of the same entity, such as **cars collection*, and those referring to multiple types within a category, such as *vehicles exhibition*. The results revealed a statistically significant preference for compounds referring to multiple types, with participants rating these compounds as more acceptable than those indicating many tokens of the same entity. Such results provide evidence that the regular plural morpheme inside compounds contributes to semantic distinctions when it denotes categorical diversity. The findings have important implications for linguistic theory, demonstrating that the regular plural morpheme within compounds is not simply a grammatical linking element observed in noun compounds of other Germanic languages, but rather one that carries morphosemantic weight. In conclusion, the study argues that the regular plural within compounds operates similarly to lexical plurals, contributing semantic distinctions beyond its function as an inflectional marker.

Index Terms—regular plural morpheme, English compounds, morphosemantic role, linking elements, lexical plurals

I. INTRODUCTION

Compounding is a word-formation process that concatenates two or more words to form a new word with a new meaning, called a compound:

- (1) a. fire + work = firework
- b. taxi + driver = taxi driver
- c. two + week + vacation = two-week vacation

The rightmost word in the compound is called the ‘head’ because it determines grammatical features such as the number and word category. For instance, as vacation in two-week vacation is a singular noun, the whole word is considered a singular noun. The left-hand constituents serve as arguments or modifiers of the head. Notably, the left member remains uninflected for number, even when a plural meaning is implied (Pinker, 1999; Pop, 2022). Thus, a two-week vacation denotes a vacation lasting two weeks, a cat lover describes someone who loves cats, and a building inspector is someone who inspects buildings. However, in some instances, well-formed compounds’ non-head elements use irregular plural forms. For example, we encounter feet-first but not *heads-first, and teeth marks but not *dentures marks, even though regular and irregular plurals are semantically matched (Sproat, 1985).

The dichotomy in behaviour between regular and irregular plurals as modifiers in compounds carries important implications for various linguistic theories, sparking long-standing debates about the structure of grammar and the functioning of morphological processes. For instance, the dual mechanism posits two processing mechanisms: one that is memory-based, where irregular forms are stored, and the other that is rule-based, where regular forms are concatenated (for a detailed discussion, see, e.g., Berent & Pinker, 2007; Clahsen, 1995; among others). The outcomes of the stored-based mechanism are allowed to feed into lexical rules such as compounding, resulting in forms such as alumni club and feet-first. However, regular plural nouns are formed in a mechanism that cannot feed into lexical rules; hence, forms such as *dentures marks and *taxis driver are ruled out.

The level ordering theory first proposed by Kiparsky (1982) also used regular/irregular plural dichotomy within compounds as a piece of evidence for the three-level-ordered mental lexicon, with level 1 for stored forms such as irregular forms, level 2 for compounding, and level 3 for regular forms. The rules of the levels apply sequentially in the sense that no lower level can be input for a higher one. In other words, irregular forms can be input for compounding whereas regular forms cannot. Again, as in dual mechanism theory, compounds that are internally inflected with a regular plural are not possible.

The connectionist theory suggests that plurals are typically excluded from compounds because English rarely has a medial plural inflection within modifier–noun sequences (Haskell et al., 2003); *smarts students, for example, does not exist. This rarity leads language users to extend this knowledge from their associative memory by not establishing a connection between the regular plural as a medial inflection and compounds. Consequently, this understanding results in

the proposal of two specific constraints: a semantic constraint that disfavors the use of plural forms in the modifier position and a phonological constraint that opposes modifiers ending with sounds characteristic of regular plurals. Because irregular plurals do not conclude with the typical sounds associated with regular plurals, language users tend to apply only a semantic constraint against their use inside compounds, thereby making them more acceptable than their regular counterparts.

However, the problem with these theories when using the regular/irregular plural dichotomy within compounds is the implicit assumption that the regular plural is strictly absent from modifiers in compounds whereas attested compounds internally inflected with a regular plural exist, as in the following example extracted from the Corpus of Contemporary American English (COCA) (Davies, 2024a):

(2) ‘But they want to see this two-step approach, because there are these conditions, perhaps, there could be an honest disagreement about what compliance with **weapons inspection** would mean.’

(3) ‘The Press and **Publications Law** and the penal code already criminalise defamation...’

Instances like (2) and (3) question the assumption of a robust absence of the regular plural inside compounds as supportive evidence for the previously mentioned theories. In level ordering theory, the ordering constraint is violated when third-level products such as *weapons* and *publications* concatenate with the nouns *inspection* and *catalogue*, respectively, at the second level. The dual mechanism constraint is violated by the involvement of rule-based products like *weapons* and *publications* with the lexical rule of compounding. For connectionism, both the phonological and semantic constraints are violated; *weapons* and *publications* are plural nouns that end with a plural-sounding phoneme.

As previously noted, some theories have overlooked this phenomenon, which would undermine their tenets. The current research argues that the absence of the regular plural inside compounds is not absolute and, thus, warrants further investigation. Two different views in literature have addressed this phenomenon. Some researchers argue that the medial regular plural in English compounds functions as a linking element, meaning that it does not affect the semantic properties of compounds (Lieber & Štekauer, 2009). Meanwhile, others suggest that the regular plural morpheme within compounds serves as a morphosemantic element (Alegre & Gordon, 1999; Al-shehri, 2014) in the sense that it adds subtle meaning and contributes to the overall semantic interpretation of the compound. The current research contributes to the existing literature by experimentally investigating the function of the regular plural morpheme inside compounds in order to provide a comprehensive understanding of this phenomenon and its implications for linguistic theory.

II. LITERATURE REVIEW

A. The Regular Plural Inside English N-N Compounds Is a Linking Element

Lieber and Štekauer (2009) argued that the presence of the regular plural morpheme in well-formed compounds should be viewed as a linking element seen in German noun–noun compounds. To understand the essence of their argument, it is essential to review the concept of linking elements in German compounds.

B. Linking Elements Inside German Compounds

A linking element is a sound unit represented as a grapheme at the junction between the components of compound words (Fenger & Harðarson, 2021; Schlücker, 2023). They function as connectors commonly observed in Germanic languages such as German, Dutch, Swedish, and Norwegian. We consider German compounds as a case of investigation. Table 1 presents the six forms of German linking elements, along with examples (Montgomery, 2001). For clarity, the linking elements are positioned between vertical bars | |.

TABLE 1
LINKING ELEMENTS IN GERMAN COMPOUNDS

Linking element	German Compound	
-en-	Frau en frage	‘women’s issue’
-es-	Kind es alter	‘childhood’
-s-	Gleichheit s prinzip	‘principle of equality’
-e-	Mäus e plage	‘plague of mice’
-n-	Seide n kleid	‘silk dress’
-er-	Kind er garten	‘kindergarten’

In many German compounds, the linking element -s-, which resembles the English plural morpheme, is argued to have originated from the genitive singular inflection -s for masculine and gender-neutral nouns (Fenger & Harðarson, 2021; Nübling & Szczepaniak, 2008; Wegener, 2008). Montgomery (2001) notes that genitive phrases in Old German underwent a long process known as lexicalisation, during which the phrasal status was gradually lost, producing what is now called genitive compounds. For instance, the genitive noun phrase *des Friedhofs^{Gen.Sg} Mauer* ‘the cemetery’s wall’ has been lexicalized. The resulting compound word is *Friedhofsmauer*, meaning ‘cemetery wall’. In this transformation, the genitive inflection -s on the first constituent of the former phrase becomes the linking element -s- in the latter constituent (Dressler et al., 2001). The lexicalisation process affected not only the form of the genitive inflection, but also its function; this form has largely lost its genitive meaning and now serves primarily as a connector between the compound’s constituents. This view is supported by various researchers (e.g., Aronoff & Fuhrhop, 2002; Neef, 2009). The semantic emptiness of the linking elements is evident in the compounds in (4) and (5):

- (4) *Liebesbrief* ‘love letter’
 (5) *Arbeitsamt* ‘employment office’

Bell (2012) points out that the linking element *-s-* inside these compounds does not correspond to any grammatical endings typically found in the base nouns. The forms *Liebes* and *Arbeits* cannot stand alone as independent words because they would be grammatically incorrect as these base nouns are feminine and the *-s-* suffix does not normally attach to feminine nouns in German. For example, the genitive forms are *Liebe* and *Arbeit*, and the plural forms are *Lieben* and *Arbeiten* (Koester et al., 2004). Therefore, as the medial *-s-* is neither a genitive nor a plural morpheme, the only plausible explanation is that it functions solely as a linking element.

Researchers have also argued that medial elements in German compounds, like *-s-*, *-er-*, and *-(e)n-*, originated from plural suffixes *-s*, *-er*, *-e*, and *-(e)n*, (Fenger & Harðarson, 2021; Wegener, 2008). The morphological/semantic function of these elements has also been debated. One argument is that the medial element in some German compounds has a morphological status due to their plural meaning; for example, *Länderspiel* refers to a match involving two countries, *Länderkunde* refers to the study of multiple nations (Dressler et al., 2001), and *Hunderennen* refers to the racing of dogs (Neef, 2009). This view, however, has been challenged by two observations. First, the non-head constituent of a German compound always has a generic interpretation, meaning it refers to a category of entities rather than specific ones. In other words, the plural interpretation of the non-head noun is understood with the presence or absence of the medial element, making it a mere connector. To illustrate, a compound with a medial element is *Autosammlung*, which refers to a collection of cars, whereas a compound without it is *Buchhändler*, which refers to a seller of books (Wegener, 2008). The more problematic observation is that there are compounds with the medial element *-er-* that impose the singular interpretation of their non-head nouns (Koester et al., 2004). Examples include *Kinderstar* ‘child star’, *Bilderrahmen* ‘picture frame’, and *Kleiderbügel* ‘coat hanger’ (Wegener, 2008), which again make the medial element appear to be a connector rather than a morphosemantic element. The question of whether the behaviour of the linking element in German compounds is similar to that of the regular plural morpheme inside English compounds is discussed in the next section.

C. Linking Elements Inside English Compounds

As in German compounds, an English compound, although very limited, may contain a linking element *-s-* that is a relic of the old genitive inflection *-es* (Bergsten, 1911; Schlücker, 2023). A limited number of compounds, particularly those ending in *-man*, include the linking element *-s-*, which is traced back to the old genitive marker *-es*. For instance, ‘kingsman’ evolved from the Old English term *caynnes mann*, and ‘townsman’ from the Middle English *tunes man* (Bergsten, 1911). This type of internal marker is no longer productive and has shed its genitive meaning.

There are also instances of English compounds with a medial element that does not originate from an old inflectional marking, but clearly reflects the plural number of the non-heads; examples include *programmes list*, *publications catalogue*, *drinks dispenser*, and *weapons inspection*. The medial plural morpheme in such examples is perceived as redundant in the sense that the plural meaning of the non-head can be inferred even in the absence of explicit plural inflection (Lieber & Štekauer, 2009). For example, both ‘programmes list’ and ‘programme list’ convey the idea of a list containing programmes; inherently, a list implies a collection of more than one item. Consider the compounds in examples (6) and (7) taken from the COCA (Davies, 2024a):

- (6) ‘Volkswagen previously said it expects its operating profit to remain flat year-on-year in 2012 despite higher **vehicles sales**.’
 (7) ‘Honda said **vehicle sales** were strong in China and in the U.S., however, thanks to the release of new models.’

Such compound pairs, where one includes an internal plural morpheme and the other does not, support the idea that this marking may be optional or non-essential, functioning as a linking element. However, this analysis does not clarify why the regular plural morpheme is optional in these compounds but considered incorrect in others (e.g., *cars sales).

D. The Regular Plural Inside English Compounds Is a Morphosemantic Element

An early but significant viewpoint from Sproat (1985) suggests that the modifier in a compound should not show plurality unless it denotes a collective group. For example, it can be understood that the regular plural in *parks department* refers to a department dealing with various categories of parks. Sproat’s view closely aligns with the heterogeneous semantics of the medial plural proposed by Alegre and Gordon (1999), who argue that using the regular plural inside compound constructions is permissible only if the plural non-head noun refers to multiple types of items (heterogeneous plurals) rather than just a single type of individual items. For example, *cars sale is considered ill-formed because the plural on the non-head implies individual cars and should therefore be unmarked for number to maintain a generic interpretation. Meanwhile, in *vehicles sale*, an attested compound, the plural non-head noun does not compromise its generic interpretation. The noun *vehicles* implies a variety of motor types—not just cars, but potentially trucks, buses, etc. Therefore, both *vehicles sale* and *vehicle sale* can refer to the sale of various types of vehicles, yet when the regular plural is used inside it, it imposes an interpretation of multiple vehicle types. Hence, Al-shehri (2014) describes the function of the medial regular plural morpheme as an ‘ambiguity resolver’ between a single type of items (homogeneous plurals) and multiple types (heterogeneous plurals).

Now consider the words *collectables* and *collectable* as modifiers of the head-noun *shop* in the following examples extracted from News on the Web (NOW) Corpus (Davies, 2024b):

(8) ‘**the collectables shop** is an opportunity to “share with other people and for people to enjoy retro items”, says owner Ben Stinson ... it will be a retro paradise filled with games, toys, computers, remote controlled vehicles, audio-visual equipment, video cassettes, posters, manages, clothes and much more including computers and games consoles for customers to play on ... There will also be some fine classic cars from the 80s and 90s, including unusual examples from Japan, Germany and France alongside retro daily driver and restoration projects.’

(9) ‘I have a small antique and **collectable shop** called Vintage Interior Pieces in the central part of Brunswick Heads.’

In example (8), the plural non-head noun ‘collectables’ emphasises that the shop sells a wide range of retrospective items that people might be interested in. Meanwhile, in example (9), the ‘collectable shop’ specialises in selling one type of collectable related to vintage interior items. Moreover, according to the NOW Corpus (more than 19.1 billion words), the term ‘collectables shop’ appears 20 times, compared to just 6 occurrences of ‘collectable shop’, which highlights the nature of such shops that typically sell a wide range of retrospective items. Examples (8) and (9) illustrate that the medial regular plural morpheme carries a semantic weight by distinguishing between what these shops sell.

To sum up, two explanations can be considered regarding the phenomenon of regular plurals within English compounds: 1) the regular plural morpheme functions as a linking element similar to that found in German compounds; and 2) the medial morpheme serves as a morphological element, resolving semantic ambiguity between a singular type and multiple types. This research aims to experimentally assess the status of these morphemes. It also seeks to contribute to the broader discourse on how morphological elements influence the semantic structure of compounds, potentially leading to a revised understanding of their function and status in linguistic theory.

III. METHODOLOGY

A. Research Question and Hypotheses

This research seeks to answer the following question: how do native speakers perceive the function of the regular plural morpheme within English compounds? Specifically, do they differentiate between cases where the morpheme retains the generic interpretation of the non-head noun by indicating multiple types and those where it violates this interpretation? To answer these questions, this study proposes two hypotheses:

Hypothesis (H1): Native speakers of English rate the regular plural morpheme that retains the generic interpretation of the non-head noun by indicating multiple types as significantly more acceptable within compounds than when it violates this interpretation.

- **Prediction 1:** If compounds that include a regular plural morpheme retaining the generic interpretation by indicating multiple types (type meaning) receive significantly higher acceptability ratings from native English speakers than those where the morpheme compromises the generic interpretation (token meaning), then it suggests that the former morpheme carries semantic weight.

Hypothesis (H0): Native speakers do not rate the regular plural morpheme that retains the generic interpretation by indicating multiple types as more acceptable than when it violates this interpretation, suggesting that the morpheme functions as a semantically empty linking element within compounds.

- **Prediction 2:** If both uses of the regular plural morpheme show no significant difference in acceptability ratings, then it will support the hypothesis that the morpheme is semantically empty and serves solely as a grammatical linking element.

B. Participants

Seventy-one native speakers of English were recruited from Prolific Academic (www.prolific.co), whose screening tool allows researcher to select participants carefully and pay them for their participation. Four participants were eliminated from the data as they responded randomly, as detected by the distractor sets. Thus, the adjusted number of participants was sixty-seven.

The reliability and transparency of this platform as an online screening tool have been reported by Palan and Schitter (2018). One essential criterion was that the participants had to be monolinguals in order to rule out the possibility that a second language might interfere with their judgements of compounds. Van Assche et al. (2009) argue that a person’s cognitive functioning and language processing are changed by a prior knowledge of a second language. Participants with no literacy difficulties were selected to read 120 sentences without problems in order to judge them without issue. All participants had at least a bachelor’s degree. Although this criterion might have limited the representation of the general population, it was considered because high educational levels are often associated with the cognitive skills needed for understanding, processing, and evaluating the subtle aspects of language usage (Pascarella & Terenzini, 2005). It was also crucial that the participants had not majored in linguistics, English, or education to ensure that their judgements were not be based on their linguistic knowledge. A questionnaire was administered before the main task to ensure the validity and reliability of the selection procedure. This questionnaire inquired about participants’ first language, second language knowledge, subject major, literacy difficulties, and educational level. All participants’ responses to the questionnaire were consistent with the initial screening criteria.

C. Materials

Four sets of sentences were created for the experiment. Set 1 included 40 sentences containing novel noun–noun compounds with a medial regular plural morpheme that denotes multiple instances of an entity, as in (10):

(10) *Contrary to expectations, the candies jar on the counter was filled with shiny red jellybeans.*

The regular plural in ‘the candies jar’ was predicted to receive a low rating as the non-head noun refers to various individual jellybeans. In this study, this set is referred to as homogeneous plurals because they refer to a collection of similar or identical items that belong to the same category.

Set 2 included 40 sentences with novel noun–noun compounds including a medial regular plural morpheme that conveys the meaning of multiple varieties within a single category, as in (11):

(11) *In landscapes management, professionals focus on maintaining diverse environments such as urban parks, coastal areas, and mountainous regions to ensure ecological balance and aesthetic appeal.*

The non-head noun in the novel compound in (11) suggests the meaning of various types of landscapes. We call this set heterogeneous plurals because they represent a range of distinct subcategories or variations within a broader category.

The contexts in which these novel compounds appear were carefully constructed to promote either a heterogeneous or homogeneous plural interpretation, which was essential because, if the compounds were presented in isolation, they could be ambiguous to participants. In such cases, participants might tend to dislike them as the typical expectation for compounds is that their non-head nouns remain unmarked for number. Hence, the current study aimed to reduce ambiguity by embedding the compounds in contexts that clearly support either a homogeneous or heterogeneous reading.

Furthermore, in both sets, we deliberately excluded phrasal compounds (e.g., ‘red candies jar,’ ‘interesting manuscripts archive’) to avoid alternative interpretations arising from their potential status as phrasal compounds (Al-shehri, 2014). In the study by Senghas et al. (1991), participants were asked to evaluate compounds like ‘modern city guide’ and ‘modern cities guide’ in contexts that could trigger either the interpretation of a city guide that is modern, represented by the bracket structure [modern [city guide]], or a guide for modern cities, represented by [[modern cities] guide]. The presence of plural internal nouns tended to support the latter interpretation, meaning a guide of modern cities, known as the recursive interpretation. Conversely, when the internal nouns were singular, participants favoured the non-recursive interpretation, meaning a city guide that is modern.

Another procedure to minimise ambiguity and misinterpretation was to exclude compounds with regularly pluralised animate non-head nouns because they might be interpreted as possessive compounds (e.g., lawyers society, students union). Such compounds can be interpreted as possessive compounds even when lacking the possessive apostrophe. Moreover, compounds with plurale tantum non-head nouns (e.g., pyjamas set, clothes rack) were also excluded to distinguish the target compounds from this special class of nouns to ensure a clear focus on the specific semantic properties under investigation (see Al-shehri, 2014) for a detailed discussion of these potential alternative interpretations).

Sets 3 and 4 included 40 ill- and well-formed sentences, respectively, that served as distractors to mask the purpose of the study and identify random answers. These distractors helped remove 4 participants from the research as they gave 6 or 7 ratings for all ill-formed sentences. All sets were randomized using the Latin square technique to guarantee that every participant was equally exposed to the different sentence types and adequately controlled for order effects and participant-specific biases.

D. Data Collection Procedures

The online experiment was built and conducted via the Gorilla platform (www.gorilla.sc), recognized for its reliability and validity (Anwyl-Irvine et al., 2020). Initially, the participants had to consent to the experiment with the assurance that they could withdraw at any time.

The experiment was built into four blocks: introduction, practice, main task, and conclusion. The first block introduced the title of the experiment, along with detailed instructions. The instructions were written in large black font on a white screen for clarity. The instructions explained in detail what was expected and how participants should respond. The participants had to rate 120 sentences on a 7-point scale, ranging from 1 as completely unacceptable to 7 as entirely acceptable. They were instructed to read each sentence without haste as there was no time constraint, and they had to rely on their intuition when making judgements. Clicking the spacebar took them to the practice block, where they had to judge five sentences. Each sentence was written in black font in the middle of a white screen. Under each sentence, a 7-point scale appeared, and the participant had to rate it before moving to the next screen. A fixation screen of 200 milliseconds appeared after each sentence. By the end of the practice session, the participants advanced to the main experiment session. When they finished, a final block appeared, announcing the end of the experiment and redirecting them to the recruitment platform, where they could receive their payments. The whole experiment took no more than 30 minutes.

IV. RESULTS AND DISCUSSION

The tendency and variability of data are summarized in Table 2.

TABLE 2
DESCRIPTIVE STATISTICS

Set	Mean	Standard Deviation
Set 1: Homogeneous plurals inside compounds	4.8	1.5
Set 2: Heterogeneous plurals inside compounds	5.7	1.3

The table shows that the acceptability ratings for sentences containing heterogeneous plurals inside compounds are higher than those for homogeneous plurals, with mean ratings of 5.7 and 4.8, respectively. This difference in mean values supports Hypothesis 1 (H1), which predicted that compounds with a medial regular plural morpheme conveying multiple types (heterogeneous plurals) would be rated more favourably by native English speakers compared to those conveying multiple instances of the same type (homogeneous plurals). The descriptive statistics reveal that the standard deviation for the ratings of heterogeneous plurals (1.3) is lower than that for homogeneous plurals (1.5). This difference indicates that participants' ratings for heterogeneous plurals were more consistent, with responses clustering more closely around the higher mean value. In other words, participants generally agreed more on the acceptability of sentences with heterogeneous plurals.

An assessment of the data distribution was conducted to ensure the validity of the findings, which is a crucial step in choosing the appropriate inferential test. As Figures 1 and 2 show, the histograms indicated a deviation from the normal distribution as we can see a skewness towards the lower end of the scale:

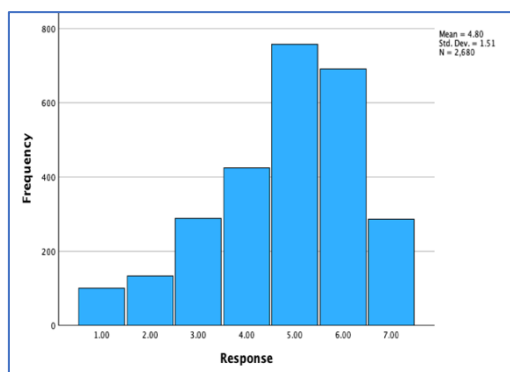


Figure 1. Mean Response to Homogeneous Plurals Inside Compounds

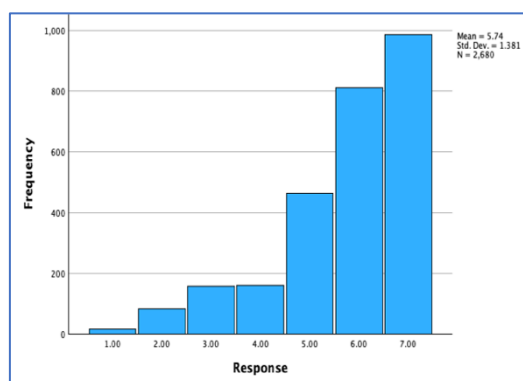


Figure 2. Mean Response to Heterogeneous Plurals Inside Compounds

Kolmogorov-Smirnov and Shapiro-Wilk tests were conducted to confirm the results statistically; both tests yielded significant results ($p < .001$). Hence, the Mann-Whitney U test was chosen as the appropriate non-parametric method for further analysis.

The dataset was also examined for outliers by calculating the Z-scores of the acceptability ratings, which is a method that can identify outliers by determining how many standard deviations a data point is from the mean, with values above 3 or below -3 being typically considered outliers (Weinberg & Abramowitz, 2008). The analysis revealed no Z-scores exceeding these thresholds; thus, the dataset is relatively homogeneous and free from extreme variations that could skew the results.

Given the non-normal distribution of the data, as confirmed by the Kolmogorov-Smirnov and Shapiro-Wilk tests, the Mann-Whitney U test was used to evaluate differences in acceptability ratings between sentences containing noun-noun compounds with homogeneous and heterogeneous regular plural morphemes. The descriptive statistics for median and range are presented in Table 3.

TABLE 3
DESCRIPTIVE STATISTICS

Sentence Type	Median	Range
Set 1: Homogeneous plurals inside compounds	5	6
Set 2: Heterogeneous plurals inside compounds	6	6

The table shows that the median acceptability rating for heterogeneous plurals was higher (median = 6) compared to homogeneous plurals (median = 5). However, both sets had an identical range of 6, indicating that the spread of participants' ratings was the same for both types of compounds. The identical range suggests that, although participants were divided in their opinions about the acceptability of plurals within compounds, this division was consistent across both homogeneous and heterogeneous plurals. In other words, there was a similar level of variability in how much participants liked or disliked both types of compounds. This wide range of ratings indicates a general discomfort or mixed feelings about the presence of plurals inside compounds. However, despite this general uncertainty, participants showed a clear preference for heterogeneous plurals, as indicated by the higher median rating.

The Mann-Whitney U test revealed a statistically significant difference between the two types of compounds ($U = 5058291.500$, $N_1 = 2680$, $N_2 = 2680$, $p = .001$, two-sided), supporting Hypothesis 1 (H1). This significant result, combined with the higher median for heterogeneous plurals, suggests that, although participants may generally dislike or feel uncertain about the use of plurals in compounds, they find compounds with internal regular plural morphemes indicating multiple types (heterogeneous plurals) to be more acceptable than those indicating multiple instances of the same type (homogeneous plurals).

These findings offer direct insights for answering the research question (i.e., How do native speakers perceive the function of the regular plural morpheme within English compounds?). The findings align with Hypothesis 1 (H1); compounds with heterogeneous plurals received higher mean and median ratings and had a lower standard deviation. The difference in the rating between heterogenous and homogeneous plurals within compounds was also statistically significant. In terms of the assumption as to whether the regular plural morpheme as a medial element in the juncture position in compounds should be considered a linking element, the findings provide evidence that it can carry morphosemantic weight when it conveys diversity within a category.

However, participants exhibited a general dislike of the regular plural morpheme inside compounds, as evidenced by the wide range of ratings for both heterogeneous and homogeneous plurals (range = 6). This result suggests that participants were equally divided in their opinions for both types, which was expected as the standard construction was a compound with a non-head noun unmarked for number. Again, the statistically different sensitivity towards both types of plurals inside compounds provide evidence that the heterogeneous plural serves a morphosemantic function rather than a connector.

Finally, as previously noted, this study faced a limitation related to the use of Mann-Whitney U test, which is not as powerful as other inferential tests (e.g., t-tests) due to the non-normally distributed data. For example, this test does not give information on effect size; it shows that the difference between heterogeneous and homogeneous regular plurals within compounds is significant, but the information on the size and meaningfulness of that difference remains indeterminate. In addition, although the sample size of 67 participants is representative, a larger sample size is suggested to help achieve a more normal distribution of data. Non-parametric tests such as t-tests can then be used to provide information on effect size.

V. CONCLUDING REMARKS

The findings of this study have key theoretical implications, particularly regarding the status of plural morphemes inside compounds as either inflectional or lexical plural. The results suggest that the heterogeneous plural inside compounds functions more as a lexical plural than an inflectional one. Lexical plurals are often characterised by their idiosyncratic interpretations because their plural meanings do not derive directly from their singular counterparts (Alexiadou, 2021). For example, the word 'goods' means 'merchandise', not 'many instances of something good'. Similarly, the interpretations of heterogeneous plurals inside compounds do not reflect the meaning of 'many instances of one type', but they go beyond the function of number marking to reflect a specialised meaning of 'multiple types of a single category'.

Acknowledging that regularly pluralised nouns are not always semantically transparent when they serve as a heterogenous function inside compounds can ease some of the challenges posed to the theories mentioned in this study's literature review section. To recap, this phenomenon threatens the ordering hypothesis of affixes by allowing a third-level inflection to interact with a second-level rule of compounding (Kiparsky, 1982). It also challenges the rule-based theory of compositional words, which prevents a compositional regular plural noun from interacting with a lexical rule such as compounding (e.g., Cunnings & Clahsen, 2007). However, when considering that the regular plural can, in some contexts, be idiosyncratic for its specialised function depending on the context, then this phenomenon does not pose a challenge to these theories, although it still challenges the phonological and semantic constraints proposed within the connectionist framework, originally suggested by Haskell et al. (2003).

Another property of lexical morphemes is the blocking effect, a well-established property in morphological constructions where the presence of one form prevents the use of another that would otherwise be grammatically correct

(Alexiadou, 2021). This effect usually happens when two forms convey similar meanings, but one is more specialised or has become conventionalized. For example, in English, the word *children* blocks the formation of *childs* as a plural for *child*, even though the regular plural rule would apply to most nouns. Similarly, the adjective *unhappy* blocks *non-happy* because ‘un-’ is the conventional morpheme used for forming the negative, even though *non-happy* would be theoretically possible under the general rule for using *non-* as a negation prefix. In these cases, the more specific or established form takes precedence and, thus, blocks the usage of the more general or rule-governed form.

The study argues that this blocking effect also occurs in the case of heterogeneous and homogeneous plurals within English compounds. A heterogeneous plural, which refers to multiple types within a category, is more specialised than a homogeneous plural, which refers to multiple tokens of the same entity. Due to its more specific meaning, the heterogeneous plural blocks the homogeneous plural in these contexts. The role of the heterogeneous plural in signalling diversity within a category offers a clearer, more semantically rich contribution than the homogeneous plural, which merely indicates a quantity of identical items. Thus, the specialised meaning carried by the heterogeneous plural blocks the use of a more general homogeneous plural, as the former is seen as more informative and acceptable in compounds.

Regardless of the theoretical model that addresses this phenomenon, the findings of the study have provided evidence that the regular plural and compounding rules interact. They reveal that the phenomenon is real and can be acceptable only if the generic interpretation of the non-head noun is not jeopardised. The native speakers’ significant acceptance of compounds with a heterogeneous regular plural on their modifiers provides evidence that the two rules can interact, but the interaction is semantically constrained. The regular plural noun should convey the meaning of heterogeneity and, hence, be tentatively encoded with the feature [REGpl, + heterogenous]. The interaction of rules is represented in the schemata in (12):

- (12) Vehicles_[REGpl, +heterogeneous] + auction_[N] → Vehicles auction
 Cars_[REGpl, -heterogeneous] + auction_[N] → *Cars auction

As *vehicles* denotes multiplicity, it is licensed to interact with another noun to form the compound word ‘vehicles auction’. *Cars*, on the other hand, denotes individual tokens; hence, it is encoded with the feature [-heterogeneous], so when it interacts with the noun *auction*, the non-head noun loses its generic interpretation.

The third implication of the findings is that they show that the phenomenon should not be polarised in the sense that compounds with an unmarked non-head noun is acceptable whereas the regular plural inside compounds is robustly unacceptable for language users. The findings, however, showed that language users can perceive the difference between a regular plural noun that retains or causes to lose the generic interpretation of the non-head of a compound. Here, the acceptability should be treated as a gradient. Hence, a tentative continuum is suggested; for the left end, we suggest a feature [unmarked_{NUM}, ± heterogenous] is acceptable whereas, for the right end, the feature [Reg_{plural}, -heterogeneous] is unacceptable, leaving in the middle of the continuum features like [Irreg_{plural}, ± heterogenous] and [Reg_{plural}, + heterogenous]. The concept of gradience is represented in (13):



Figure 3. Continuum of Plural Modifier Acceptability in Compounds

The most acceptable forms of and more productive compounds are undeniably those whose modifiers are unmarked for number (e.g., *car collection*) whereas the least acceptable compounds are those whose modifiers are regularly pluralised, denoting many tokens for compromising the generic interpretation of their non-head nouns (e.g., **cars collection*). In the middle, near the most acceptable forms of compound, we find compounds with irregular plural modifiers. According to the literature, attested compounds exist, and they are seen as somewhat acceptable for idiosyncratic reasons (e.g., *teeth marks*) (Sproat, 1985; Berent & Pinker, 2007). Along the continuum, we see compounds with modifiers carrying the features of regular plurality and heterogeneity, which proved to be more acceptable by the participants in the present research than the forms on the right end of the continuum (e.g., *vehicles sales*).

To conclude, this study provides crucial insights into the semantic role of regular plural morphemes within English noun–noun compounds, particularly emphasizing the distinction between homogeneous and heterogeneous plurals. The results show a clear preference for heterogeneous plurals, where plural morphemes indicate multiple types within a category, suggesting that regular plural morphemes carry significant morphosemantic weight. This finding challenges the notion of regular plurals as mere grammatical linking elements, demonstrating instead that they play an active role in contributing to the meaning of the compound, especially in cases where they distinguish between different types of entities.

The findings support the view that regular plurals in compounds behave similarly to lexical plurals, adding subtle meaning to the overall construction rather than functioning purely as inflectional markers. This finding has significant implications for existing linguistic theories, as it highlights the need to reconsider how regular plural morphemes interact with compounding. In light of these findings, future research could explore how plural morphemes operate across different contexts and languages, contributing to a more refined understanding of the interface between morphology and semantics in compound formation.

REFERENCES

- [1] Alegre, M., & Gordon, P. (1999). *Why compounds researchers aren't rats eaters: Semantic constraints on regular plurals inside compounds*. Columbia University.
- [2] Alexiadou, A. (2021). Lexical plurals. In P. C. Hofherr & J. Doetjes (Eds.), *The Oxford handbook of grammatical number* (pp. 242-256). Oxford University Press.
- [3] Al-Shehri, A. A. (2014). *Regular plural inside English compounds within the theory of base-driven stratification* [Unpublished PhD thesis]. University of Edinburgh.
- [4] Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior research methods*, 52, 388-407. <https://doi.org/10.3758/s13428-019-01237-x>
- [5] Aronoff, M., & Fuhrhop, N. (2002). Restricting suffix combinations in German and English: Closing suffixes and the monosuffix constraint. *Natural Language & Linguistic Theory*, 20(3), 451-490. <https://doi.org/10.1023/A:1015858920912>
- [6] Bell, M. (2012). *The English noun-noun construct: A morphological and syntactic object*. *Mediterranean Morphology Meetings*, 8, 59-90.
- [7] Berent, I., & Pinker, S. (2007). The dislike of regular plurals in compounds: Phonological familiarity or morphological constraint? *The Mental Lexicon*, 2(2), 129-181. <https://doi.org/10.1075/ml.2.2.03ber>
- [8] Bergsten, N. (1911). *A study of compound substantives in English*. Almqvist & Wiksell.
- [9] Clahsen, H. (1995). German plurals in adult second language development. In L. Eubank, L. Selinker, & S. Gass (Eds.), *The current state of interlanguage: Studies in honour of William E. Rutherford* (pp. 123-137). John Benjamins.
- [10] Cunnings, I., & Clahsen, H. (2007). The time-course of morphological constraints: Evidence from eye-movements during reading. *Cognition*, 104(3), 476-494. <https://doi.org/10.1016/j.cognition.2006.07.010>
- [11] Davies, M. (2024a). *The corpus of contemporary American English (COCA)*. Retrieved March 4, 2024, from <https://www.english-corpora.org/coca/>
- [12] Davies, M. (2024b). *The NOW corpus (News on the Web Corpus)*. Retrieved March 4, 2024, from <https://www.english-corpora.org/now/>
- [13] Dressler, W., Libben, G., Stark, J., Pons, C., & Jarema, G. (2001). The processing of interfixed German compounds. In G. Booij & J. van Marie (Eds.), *Yearbook of Morphology* (pp. 185-220). Kluwer.
- [14] Fenger, P., & Harðarson, G. R. (2021). *The plurals are not what they seem: On linking morphemes and inflection in West Germanic*. Retrieved May 10, 2024, from https://gislihardarson.wordpress.com/wp-content/uploads/2021/09/l_articl_-4.pdf
- [15] Haskell, T., MacDonald, M., & Seidenberg, M. (2003). Language learning and innateness: Some implications of compounds research. *Cognitive Psychology*, 47, 119-163. [https://doi.org/10.1016/S0010-0285\(03\)00007-0](https://doi.org/10.1016/S0010-0285(03)00007-0)
- [16] Kiparsky, P. (1982). Lexical phonology and morphology. In I. Yang (Ed.), *Linguistics in the Morning Calm* (pp. 3-91). Hanshin.
- [17] Koester, D., Gunter, T., Wagner, S., & Friederici, A. (2004). Morphosyntax, prosody, and linking elements: The auditory processing of German nominal compounds. *Journal of Cognitive Neuroscience*, 16(9), 1647-1668. <https://doi.org/10.1162/0898929042568541>
- [18] Lieber, R., & Štekauer, P. (Eds.). (2009). *The Oxford handbook of compounding*. Oxford University Press.
- [19] Montgomery, B. (2001). *Studies in German and English morphology with special reference to 'linking elements'* [Unpublished PhD thesis]. University of Edinburgh.
- [20] Neef, M. (2009). IE, Germanic: German. In R. Lieber & P. Štekauer (Eds.), *The Oxford handbook of compounding* (pp. 386-399). Oxford University Press.
- [21] Nübling, D., & Szczepaniak, R. (2008). On the way from morphology to phonology: German linking elements and the role of the phonological word. *Morphology*, 18, 1-25. <https://doi.org/10.1007/s11525-008-9120-7>
- [22] Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, 22-27. <https://doi.org/10.1016/j.jbef.2017.12.004>
- [23] Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: a third decade of research* (2nd edn). Jossey-Bass.
- [24] Pinker, S. (1999). *Words and rules: The ingredients of language*. Weidenfeld & Nicolson.
- [25] Pop, I. B. (2022). On some idiosyncrasies exhibited by the plural of compounds. *Meridian Critic*, 41, 265-276.
- [26] Schlücker, B. (2023). *Compounding and linking elements in Germanic*. Retrieved February 9, 2024, from <https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-954>
- [27] Senghas, A., Kim, J., Pinker, S., & Collins, C. (1991). Plurals-inside compounds: Morphological constraints and their implications for acquisition. In *sixteenth annual Boston university conference on language development* October (pp.18-20).
- [28] Sproat, R. (1985). *On deriving the lexicon*. MIT.
- [29] Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science*, 20(8), 923-927. <https://doi.org/10.1111/j.1467-9280.2009.02389.x>
- [30] Wegener, H. (2008). The grammaticalization of linking elements in German. In S. Seoane & M. López-Couso (Eds.), *Theoretical and empirical issues in grammaticalization* (pp. 333-354). Benjamins.
- [31] Weinberg, S. L., Abramowitz, S. K. (2008). *Statistics using SPSS: An integrative approach*. Cambridge University Press.

Amira Abdullah Alshehri is an Assistant Professor of Linguistics at the College of Education and Arts, University of Tabuk, Saudi Arabia. She earned her PhD from the University of Edinburgh. Her research interests include lexical processing, generative morphology, corpus linguistics for specialized fields such as medical terminology, English for Specific Purposes (ESP), and second language acquisition. Email: aa.alshehri@ut.edu.sa ORCID: <https://orcid.org/0000-0002-5555-0937>