

# The Function of Reduplicative Agreement in the Acquisition of Noun Classes

## Abstract

Noun classification systems are common in languages across the world, in which nouns are categorized into subclasses of shared features. These shared features may be semantic, phonological or morphological characteristics of the noun, but noun class is primarily determined by patterns of grammatical agreement outside of the noun. While previous research suggests that such systems carry a communicative advantage in language processing, they seem to be costly in terms of acquisition - speakers need to learn both the subclass of each noun and the corresponding grammatical agreement. However, research into language acquisition in Sesotho by Demuth and Weschler (2012) suggests that agreement morphology may be an important feature in noun class learning as it provides learners with multiple cues to class. This study therefore offers a first direct test of the role of morphological agreement in the acquisition of noun classes. By comparing two common types of morphological agreement in an artificial language learning experiment - reduplicative agreement and arbitrary agreement - we present evidence that reduplicative morphological agreement facilitates the acquisition of noun class systems.

## Introduction

Noun classification systems are common in languages across the world, in which nouns are categorized into subclasses of shared features. Such systems can be grouped into three main types: grammatical gender, noun class and noun classifier systems. Examples of each type are given in (1) - (3) below and, for the purpose of this discussion, will serve as a reference of these systems throughout the paper. The first language, German (1), has a gender system where nouns fall into three possible classes - female, male or neuter - expressed on articles (*die/der/das*) and modifiers preceding the noun.

### (1) German

|                      |                      |                    |
|----------------------|----------------------|--------------------|
| <i>a. die Person</i> | <i>b. der Mensch</i> | <i>c. das Kind</i> |
| the.F person         | the.M human          | the.N child        |
| 'the person'         | 'the human'          | 'the child'        |

By comparison, Oshikwanyama (2), a Bantu language spoken in Namibia and Angola, categorises nouns into 9 classes. As is characteristic for Bantu languages, noun class is expressed through prefixes on the noun, modifiers and on verbs. This can be seen in the *omu-* prefix on the adjective and noun referring to 'woman' in (2) b. and the *oku-* prefix on the verb. Both these prefixes reveal that the noun 'woman' belongs to Class 1, while the *(o)shi-* prefix on the noun for goat and the numeral shows that 'goat' falls under Class 5. By comparison, both nouns in (2) b.

belong to Class 7 and therefore receive the same morphological marking (*ou-* on the noun, *va-* on the numeral).

(2) Oshikwanyama

- a. *Omu-kulupe omu-kulukadi oku-na oshi-kombo shi-mwe.*  
 1.SG-old 1.SG-woman 1.SG-have 5.SG-goat 5.-one  
 ‘The old woman has one goat.’

- b. *Ou-matyona va-tatu ove-na ou-mbwa va-vali.*  
 7.PL-boys 7.PL-three 7.PL-have 7.PL-dog 7.PL-two  
 ‘The three boys have two dogs.’

Finally, (3) displays examples of numeral classifiers in Cantonese Chinese. Depending on semantic features of the noun, a large variety of the classifier particles are used obligatorily for any definite or numerical noun phrase. For example, the classifier *go[3]* is used with nouns referring to humans, *zek[3]* is used for animals and *zi[1]* is used to refer to nouns denoting long, narrow objects (Tsang & Chambers, 2011).

(3) Cantonese Chinese

- |                               |                                |                               |
|-------------------------------|--------------------------------|-------------------------------|
| a. <i>sam[1] go[3] jan[4]</i> | b. <i>sam[1] zek[3] gau[2]</i> | c. <i>sam[1] zi[1] bat[1]</i> |
| Three cl people               | three cl dogs                  | three cl pens                 |
| ‘Three people’                | ‘Three dogs’                   | ‘Three pens’                  |

(Culbertson, Jarvinen, Haggarty & Smith, 2018)

When comparing the above examples, differences between these three types of classification systems become clear on both the semantic and the formal level. Firstly, classifier systems are not semantically arbitrary (Senft, 2000), i.e. the use of the classifier is determined by the meaning of the noun. By contrast, gender and noun classes vary greatly in their semantic reliability. While some gender languages do follow overall patterns of sex in assigning grammatical gender (e.g. ), the examples of German above show that grammatical gender here does not necessarily correspond to sex. Instead, gender assignment appears entirely arbitrary: all three nouns denote human beings with no references to biological gender, yet each noun falls under a different grammatical gender. Bantu languages are similarly semantically unreliable in their assignment of noun classes. The most reliable semantic category is Class 1 - including most nouns for human beings. Other noun classes include a mix of semantic categories such as animals, plants, inanimates, size differences, and abstract nouns. Many researchers argue that these meaning distinction originally may have been part of an even more elaborate system of semantically distinct noun classes in Proto-Bantu (Demuth 2000), but in Oshikwanyama nouns

like ‘boy’ in (2) b. fall into the same class as ‘dog’. This is because Class 7 includes nouns referring to anything of a small size, but also several animals. Hence, this example demonstrates that even the most reliable semantic feature [+human] in Oshikwanyama may be overridden by other semantic features such as [+small] in determining a noun’s class. Consequently, this suggests that in most noun class languages, semantics alone cannot be sufficient in indicating noun class for a learner acquiring the grammatical categories of the nouns.

Secondly, the examples above demonstrate differences in the formal expression of the class in noun classification systems. Classifier languages usually have only one cue for classification (Grinevald, 2000, p. 62), e.g. in Cantonese the classifier comes in the form of the morpheme preceding the noun. By contrast, gender and noun class languages are characterized by grammatical agreement (Corbett, 1991). Expression of this agreement can be in the form of class markers on the noun, other constituents of the noun phrase, and extends to the expression of class markers outside of the noun phrase (as seen on verbs in Oshikwanyama). Additionally, agreement of noun class can be reduplicative, i.e. it is expressed with the same phoneme or morpheme across nouns and their modifiers, as in (2) a. (*‘omu-’* on the adjective and noun for ‘woman’/*‘(o)shi-’* on the noun and numeral for ‘goat’). This seems to be a particularly efficient manner of agreement, as the same form is re-used across sentence constituents in expressing the same meaning, i.e. the noun class. Alternatively, agreement can be arbitrary as with the prefixes in (2) b. (*‘ou-’*-noun co-occurring with *‘va-’*-numeral). The learning implications for both types of agreement are discussed in more detail below. In sum, noun class languages (including gender) thus differ from classifier languages in their semantic reliability and their use of grammatical agreement. Furthermore, the patterns of agreement amongst noun class languages may differ between being reduplicative and/or arbitrary.

Despite these differences in semantic and formal features, all noun classification systems share features of apparent redundancy. Several researchers have questioned the function of numeral classifiers specifying parts of a noun’s meaning when this information is expressed in the noun itself (Denny, 1976; Grinevald, 2000). In other words, why does the classifier in (3) specify numeration of humans, animals or long, narrow objects when these specifications are already contained in the nouns’ semantic features? Similar questions can be asked for noun class languages. Due to the lack of complete semantic reliability, noun class cannot be deduced based on the meaning of the noun alone, which makes the noun class markers useful in cuing the grammatical category (e.g. articles indicating gender in German or prefixes cuing noun class in Oshikwanyama). However, despite the grammatical necessity of these markers, the patterns of agreement extending beyond the noun phrase seem to create another type of redundancy in form. The example of Oshikwanyama (2) b. above displays three morphological markers (*ou-*, *va-* and *ove-*) each expressing the same meaning, which is the noun class of the subject, ‘boy’. Thus, all

noun classification systems seem to reflect degrees of morphological redundancy by expressing meanings already conveyed elsewhere.

One would expect to find explanations for such apparent structural redundancy in looking at the function of noun class agreement in communication or language acquisition. Several researchers, such as Kirby, Tamariz, Cornish and Smith (2015), argue that the origins of linguistic structures lie in competing pressures acting on languages during communication or acquisition. According to their findings, learning favours language simplicity, causing languages to become more compressible. Contrastively, pressures for expressivity or ambiguity avoidance are introduced during language use, thus decreasing compressibility. These two constant pressures for simplicity and expressivity during acquisition and use respectively, are therefore believed to result in an overall pressure for efficiency in languages. If this view of language should be accurate, then the apparent structural redundancy of noun class languages should serve a function in language use or acquisition. In other words, if noun class agreement should introduce a benefit during communication that is high enough to counteract pressures for morphologically simpler systems, one would expect such a system to be introduced and maintained in the use of language. Likewise, if some structures of noun class systems are simpler than others, while maintaining the same degree of expressivity, then one would expect to find these in languages as a result of generations of language learners.

An example of such structural efficiency would be reduplicative morphology, which is a widespread feature in languages across the world and seems to achieve a perfect trade-off between simplicity and expressivity. While learners are tasked with the learning of only one form per meaning, this form can be re-used multiple times to increase its expressive potential. Cabellero & Kapatsinski (2012) show that such repeated use of morphological markers expressing a single meaning in Choguita Rarámuri can play a key role in achieving communicative clarity. As mentioned above when discussing the example of Oshikwanyama (2) this form of morphological agreement is a common feature in noun class languages and may therefore also play a key role in meeting pressures for expressivity and simplicity in these languages.

Multiple studies have investigated the function of noun class marking in communication and present evidence for its role in language processing (Corbett, 1991; Zubin & Köpke, 2007, 1981; Aylett & Turk, 2004; Gahl & Garnsey, 2006; Manin, 2006; Levy, 2008; Tsang & Chambers, 2011;). Tsang, Chambers & Mozuraitis (2017), for example, suggest that classifiers preceding the noun provide an important aid in word identification in Cantonese. Thus, the high amount of compounds in the language presents a processing challenge in running speech. Additionally, the stress-timed nature of Cantonese prevents speakers from using durational contrasts in differentiating compounds from isolated words. As their results show, these challenges are

overcome by syntactic cues such as noun classifiers that cue the set of potential subsequent nouns. Similarly, Dye et al. (2017) take an information theoretic approach to show that gender in German, as expressed on articles and adjectives preceding the noun, can play an important role in information processing. They find that exceptions to gender assignment amongst semantically related nouns in German correlates with the highest frequency of use. In other words, nouns that are used the most, tend to belong to a different gender than semantically related ones. This means that similar to classifiers in Cantonese, gender markers preceding the noun in German greatly reduce the set of possible noun candidates, thus facilitating sentence processing. Altogether, such studies suggest that the complexity of noun class systems may indeed be a function of language processing and communication.

Finally, this raises the question that is of interest for this study - how do children acquire such complex systems? Agreement patterns of noun class systems require the tracking of long distance dependencies, a cognitive skill that is shown to develop only after around 14 months (Culbertson, Koulaguina, Gonzalez-Gomez, Legendre, Nazzi, 2016). Additionally, since noun class languages frequently rely on a mixture of semantic and morphophonological cues in determining the noun class, learners are presented with a complex problem in tracking both deterministic and probabilistic cues across domains, including exceptions that need to be memorized (Culbertson, Smith, Jarvinen, Haggarty, 2018a; Arnon & Ramscar, 2012; Braine, 1987; Frigo & McDonald, 1998; Levy, 1988; Taraban, 2004). Nevertheless, acquisition research of various gender and noun class languages such as Icelandic, Germanic, and Romance languages shows that children learn these categories and their grammatical agreement patterns by the age of 3 or earlier (see Corbett, 1991, 2006 for reviews). In a study of Polish-speaking children, Smoczynska (1986) shows that grammatical gender is mastered as early as age 2. This is one of the earliest documented ages by which gender seems to be acquired and Smoczynska suggests that this may be due to the high phonological transparency of gender marking in Polish.

Several studies have in fact found that learners predominantly rely on phonological rather than semantic cues when acquiring gender or noun classes (Culbertson, Jarvinen, Haggarty, Smith, 2018b; Levy, 1981; Mills, 1985; Perez-Pereira, 1991; Rodina-Westergaard, 2012; Gagliardi & Lidz, 2014). In a seminal paper, Karmiloff-Smith (1979) found that French-speaking children as old as 9, privilege phonological cues over semantic cues in determining the noun class of novel nouns. A similar result is presented by Frigo and McDonald (1998). In an artificial language learning experiment, participants were taught pairs of greetings and names. They were told that the names differed in terms of the person's social status to which they belonged. Accordingly, one pair of greetings was used for people of a higher status and the other two greetings occurred with a person of lower social status. In one test group, 60 % of the names in each subclass were phonologically marked, while 40 % of the names were unmarked. As expected, phonologically marked names were learnt far better than phonologically unmarked ones.

This prioritizing of formal cues during acquisition would suggest that the elaborate morphological agreement system of Bantu languages provides learners with a particularly helpful framework for learning the noun classes. As the example (2) of Oshikwanyama shows above, noun class prefixes in Bantu languages create a high degree of morphological transparency, regularity and pervasiveness by marking class on every word in the sentence. Demuth & Weschler (2012) present evidence on the acquisition stages of 2-3 year olds learning Sesotho, a Southern Bantu language. Their study shows that children, by the age of 2;1 already performed near or above 90% accuracy in their correct use of noun class agreement. The vast majority of the few recorded mistakes in their study happened for nouns where null-prefixes are permitted in the language. In other words, for words where children had less exposure to the noun prefix, they performed worse in allocating it to the right noun class. While Demuth's and Weschler's (2012) research only presents data from 3 Sesotho-speaking children, it is thus worth investigating whether agreement patterns as they are found in Bantu languages serve as an aid in the acquisition of noun classes.

The current study therefore aimed to investigate the function of agreement morphology in learning noun class systems. For this purpose, an artificial language learning experiment was designed comparing three morphological conditions and their usefulness in helping participants learn noun subclasses. Experimental conditions varied in their type and degree of morphological agreement. One group of participants was taught a noun class system with no agreement, but only one class marker on the noun itself. Such a system would correspond to classifier languages in that only one cue is available for the learning of noun category. The other two experimental conditions mirrored agreement morphology as it is found in Bantu and other noun class languages. In one condition, participants learned an artificial language with reduplicative morphology similar to the agreement in (2) a. In another condition, participants were taught a language with arbitrary agreement akin to agreement patterns in (2) b. Examples of these three conditions are summarized in (4) below:

|     |                          |                            |                                |
|-----|--------------------------|----------------------------|--------------------------------|
| (4) | No Agreement:            | <i>cleep-ki</i><br>noun-CL | <i>grome</i><br>modifier       |
|     | Reduplicative Agreement: | <i>cleep-ki</i><br>noun-CL | <i>grome-ki</i><br>modifier-CL |
|     | Arbitrary Agreement:     | <i>cleep-ki</i><br>noun-CL | <i>grome-pa</i><br>modifier-CL |

These three conditions enable a clear comparison of learning cost and benefit with varying types of agreement morphology. With reduplicative agreement, learners have multiple cues for the noun class and only need to learn one marker per subclass to create the agreement in the

language. With arbitrary agreement, learners have multiple cues for the noun class available, but need to learn more than one marker for agreement. Finally, in the no agreement condition, learners only have one cue available that is giving away the category of the noun, but also do not have the cost of learning more than one marker per class. This comparison is summarized in Table 1:

| Condition               | Multiple Cues for Class | 1 Marker per Class |
|-------------------------|-------------------------|--------------------|
| Reduplicative Agreement | ✓                       | ✓                  |
| Arbitrary Agreement     | ✓                       | ✗                  |
| No Agreement            | ✗                       | ✓                  |

Table 1: Comparing benefits and costs of varying agreement conditions in learning noun classes.

Based on the above comparison, we hypothesised that morphological agreement between noun class markers is better than no agreement in the acquisition of noun subclasses due to the increased availability of noun class cues. Specifically, we suggest that reduplicative agreement is the most efficient type of morphological agreement for learning noun class systems. In our experiment, this condition is therefore predicted to yield the best results for the acquisition of noun class and learning the correct agreement between markers. By contrast, we hypothesised that arbitrary agreement is still helpful in terms of offering multiple cues for noun class, but due to the cost of learning the arbitrary agreement between markers, participants in this condition are expected to perform worse both in learning the categories of nouns as well as the agreement between markers. Finally, no agreement is hypothesised to be the least helpful in learning noun subclasses, since participants are exposed to fewer cues indicating the noun class. This last condition is therefore expected to yield the lowest results in the acquisition of noun class languages.

## *Methodology*

The experiment was modelled after a miniature artificial language learning experiment used by Culbertson, Gagliardi & Smith (2016). Participants in this experiment go through an exposure phase and several testing phases. During exposure, they see pictures and hear corresponding descriptions in the artificial language. In the testing phases, participants are then shown previously seen pictures and asked to provide the corresponding descriptions.

### *Participants*

Participants were 60 native English-speaking adults recruited at the University of Edinburgh, aged 19 - 31 (mean age = 24 years). Participants were paid £ 5, and the experiment lasted between 25 – 30 minutes.

### *Materials*

Visual stimuli consisted of photos of humans, animals, and inanimate objects. The visual stimuli were allocated to two noun classes, each with 15 members in total, to create a partially reliable [+human] class and a partially reliable [+animal] class. Similar to the cue reliability in Frigo and McDonald (1998), 60% of nouns in each class exhibited the relevant semantic feature, i.e. [+human] or [+animal], while the remaining 40% of nouns in each class were inanimate objects. Each picture was paired with a nonce noun, therefore the language consisted of 30 nouns total. In addition to this, there were 2 nonce post-nominal number words corresponding to 2 and 3, and a set of class markers (2 or 4, depending on the condition). Nouns and numerals were drawn at random from 50 CVC monosyllabic nonce words, which were taken over from a previous noun class learning experiment by Culbertson, Gagliardi and Smith (2016). Noun class markers consisted of the two suffix-pairs (-mu [mu]- and -to [tɔ]; -ki [kɪ:] and -pa [pɑ]). Depending on the experiment condition (see below), these suffixes were attached to nouns only, or to nouns and numerals. Experiment conditions also determined whether only two or all four markers were used. Words were presented auditorily and orthographically. Below is a summary of the language system:

- 30 Nouns: 9 human nouns, 9 animal nouns, 12 inanimate object nouns;
- 2 post-nominal numerals (list the forms), meaning two and three.
- 4 suffixes: -mu, -to, -ki, -pa (used depending on condition, see below).

As described above, the partial semantic reliability of the noun class categories resembles noun subclasses in many natural languages, including Bantu languages. It also resembles the proportion of cue reliability in Frigo and McDonald (1989), which should ensure that participants will be able to learn the language despite partially conflicting semantic cues. However, the categorically unreliable inanimate nouns are also crucial in understanding acquisition differences between the morphological conditions. Participants can rely on two cues in learning the correct noun class of animate nouns: semantics (+human/+animal) and morphology (noun class suffix). For inanimate nouns, morphology is the only information revealing the noun class category and therefore offers the most direct comparison between learning success of the different morphological conditions.

### *Design*

Participants were divided into three conditions which differed on two dimensions: the presence or absence of class-based agreement and whether there was reduplication of class markers. In the reduplicative agreement condition, participants were exposed to a language with reduplicative class-based agreement; one of two matching class suffixes was present on both the noun *and* the numeral. In the arbitrary agreement condition the form of the suffix on the noun did not match the form of the suffix on the numeral. Finally, participants in the no agreement condition were exposed to a language without class-based agreement; one of two class suffixes was present on the noun, indicating the noun class, but no agreement was triggered on numerals.

| Condition/Class | [Human] | [Animal] |
|-----------------|---------|----------|
|-----------------|---------|----------|

|                         |                      |                      |
|-------------------------|----------------------|----------------------|
| Reduplicative Agreement | noun-ki +numeral-ki  | noun-mu +numeral-mu  |
| Arbitrary Agreement     | noun-ki + numeral-pa | noun-mu + numeral-pa |
| No Agreement            | noun-ki + numeral    | noun-mu +numeral     |

Table 2. Example input structure for each of the three experiment conditions which vary based on the presence or absence of a redundant cue to class (agreement on the numeral) and whether the redundant cue was reduplicative.

### *Procedure*

Participants were told that they would learn a small part of a foreign language. Each participant was seated in an individual experiment booth. Instructions for each stage of the experiment were provided on screen. Participants wore headphones for the entirety of the experiment and their responses were recorded through a microphone.

The experiment consisted of three stages; one training stage and two testing stages. During training, participants saw pictures and heard a description including the noun, the suffix and, if more than one object was displayed, the numeral (with a suffix in the reduplicative agreement and arbitrary agreement conditions). They were asked to repeat each word or phrase that they heard, before continuing to the next one. In addition, the words and phrases were presented orthographically below the images (see Figure I below). The training stage consisted of 180 trials; participants saw each noun once on its own and with each of the possible numbers (two and three). Training lasted around 15 minutes.



Figure I: Examples of screens participants saw during exposure in the reduplicative agreement condition. The image on the left shows a singular trial, i.e. participants only saw one image and only heard the noun

with the noun class suffix. The image on the right shows a plural trial where participants saw two/three images and heard the noun plus the numeral word (as well as the noun class suffixes on both words).

During the first stage of testing, participants were tested on singular nouns (without the modifying numerals). Here, participants saw an image they had previously seen during training and heard the corresponding noun. However, the suffix was replaced by a ‘beep’. They were then asked to provide the correct suffix. This stage assessed whether participants learned the category markers for the noun subclasses in each condition and included 60 trials (two for each noun). Note that in contrast to Frigo & McDonald (1998), no generalization trials (i.e., novel objects) were included during testing since the noun class categories are only partially reliable, i.e. the correct class for novel inanimate nouns is impossible to determine.

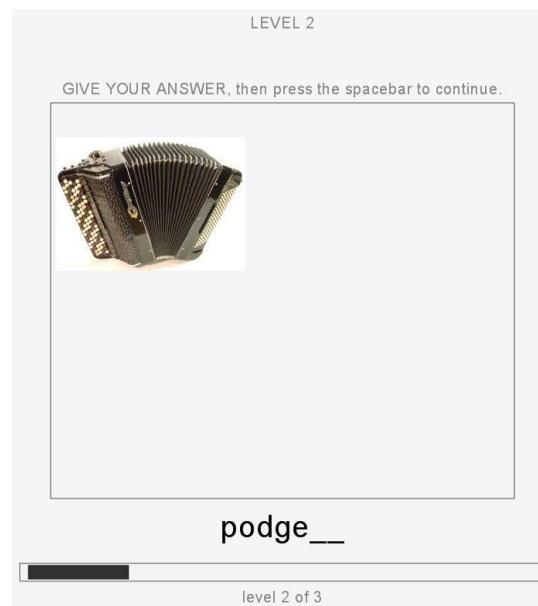


Figure 2: Example of a screen during the noun class test where participants were shown single images and a noun. They were then asked to complete the word providing the right marker.

During the second stage of testing, participants were tested on their knowledge of the agreement properties in the language. Participants saw either two or three duplicates of the same image, and heard the corresponding noun and numeral without the suffix, e.g. noun-*beep* + numeral-*beep*. They then were asked to produce the phrase including the correct suffix(es). This testing stage also consisted of 60 trials (each noun with both numbers). While the singular testing stage assessed learning of the noun class as expressed by the noun class marker, the plural testing stage revealed whether participants learned the agreement between the markers in the reduplicative and the arbitrary agreement condition (participants in the no agreement condition also went through both testing stages, but data from this condition was excluded for tests of agreement). Consequently, these two tests will be referred to as the noun class test and the agreement test, respectively.

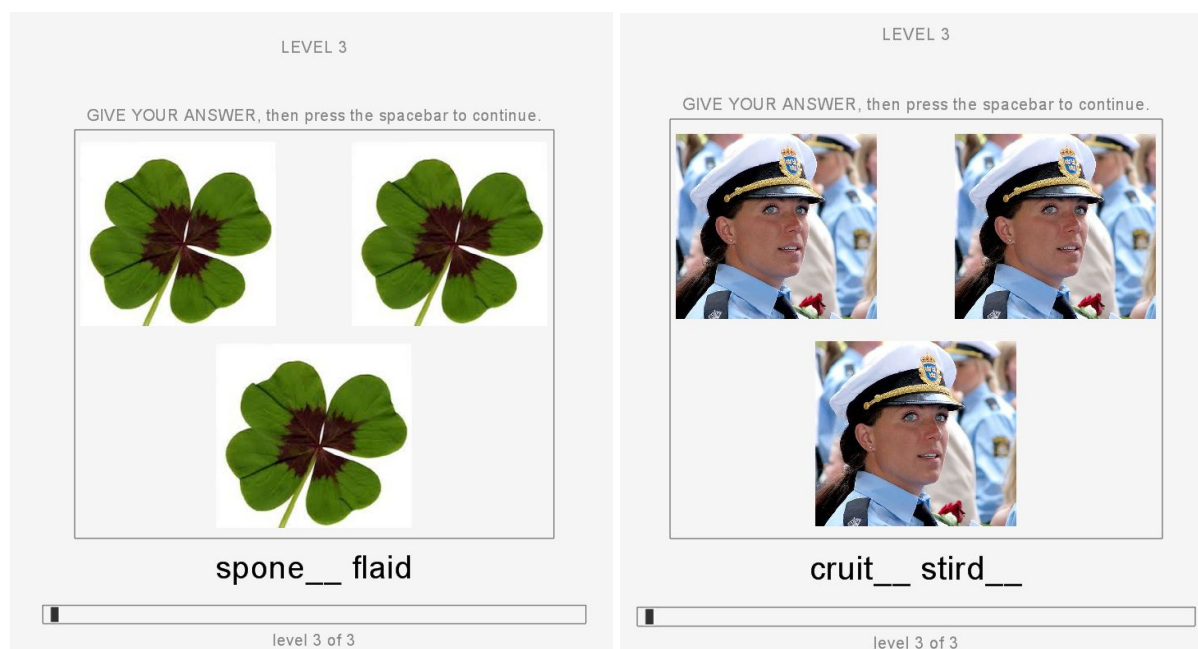


Figure 3: Examples of screens seen during the agreement test in the no agreement condition (left) and either of the agreement conditions (right). Participants were shown two or three images and asked to complete the phrases.

### *Transcription*

All responses were subsequently transcribed by the author. Transcription was limited to the noun class markers as these were the point of interest for both testing phases. In other words, even though participants in the noun class test responded with both the noun and the class marker, only the marker for each trial was transcribed. Similarly, while participants responded with noun, numerals and markers in the agreement test, only markers were transcribed. If participants failed to respond to a trial, (which happened in only 2 of 7200 trials), the trial was marked as not applicable ('NA'). For ambiguous responses, i.e. where one or more phonemes differed from the target form, the researcher transcribed it as it was heard. For example, when participants responded [pə] instead of [pa] or [tu] instead of [tɔ], the heard form was written down. Finally, each trial was given a binary outcome score - '1' if the target form and the response matched, '0' if they did not.

## Results

This study made the following predictions: reduplicative agreement would yield better learning results of noun class than no agreement or arbitrary agreement. We also predicted that participants in the no agreement condition would perform worse than the other two conditions as they were dealing with only one category cue on the noun. Hence, results of the three conditions in participants' ability to learn the noun class categories with the corresponding marker should be ranked in the following way:

1. Reduplicative Agreement
2. Arbitrary Agreement

### 3. No Agreement

As explained above, inanimate trials were of particular interest in revealing acquisition differences between the three conditions. While animate trials provided participants with two sources of information about noun class category - semantics and morphology - inanimate trials only offered a morphological cue about the noun class. Hence, we expected better results for learning of animate nouns than inanimate nouns across all conditions, but, importantly, we predicted to see smaller discrepancies between animate and inanimate trials in the morphological condition that is thought to be most conducive to noun class learning. In other words, reduplicative agreement was predicted to show the smallest difference between participants' success in category learning of animate and inanimate nouns.

Secondly, we predicted that participants in reduplicative agreement would perform better at learning the agreement between the markers than participants in arbitrary agreement. Consequently, success in learning the correct agreement should be ranked in the following way:

1. Reduplicative Agreement
2. Arbitrary Agreement

#### *Noun Class Learning*

To test our first set of predictions, we first look at the correct use of the noun suffix in the noun class test. The noun suffix reflects whether participants learned the correct category for each noun. Across all three conditions, participants had an accuracy rate of 83% in providing the right noun suffix during the noun class testing stage and thus performed well above chance in learning the noun subclasses. Figure 4 below shows the performance of participants across all three conditions.

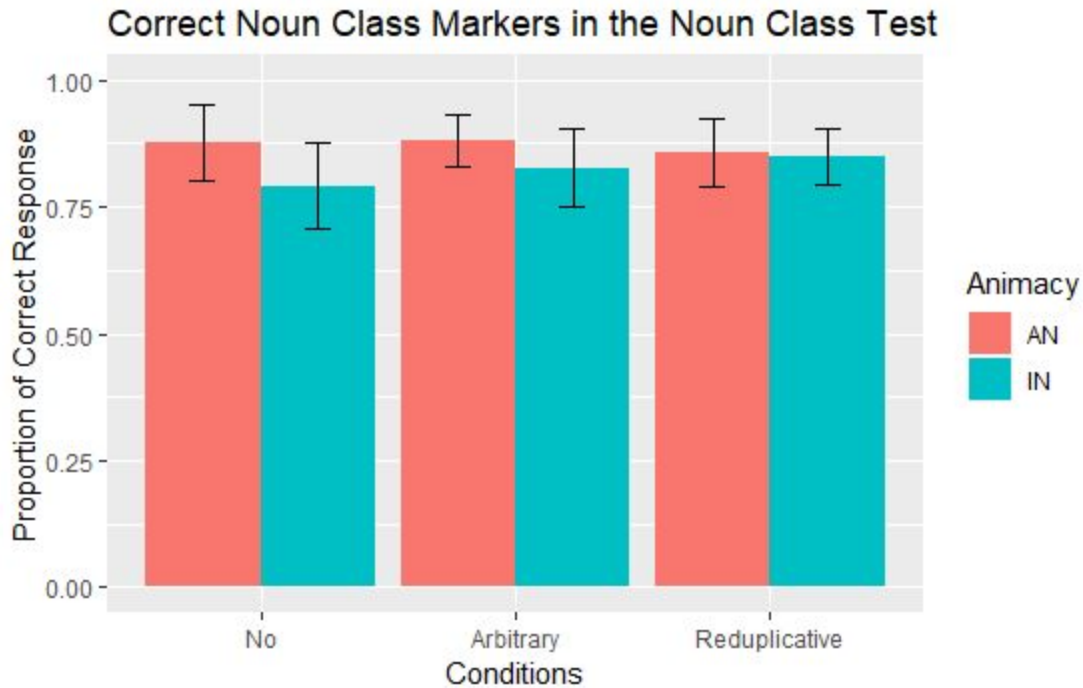


Figure 4: A comparison of participant performance in the three morphological conditions (No Agreement, Arbitrary Agreement and Reduplicative Agreement) in the noun class test. Red bars show the proportion of correct responses for animate trials, blue bars indicate correct responses for inanimate trial. Error bars show a 95% confidence interval for each condition.

Across all conditions, participants seem to have learned the subclasses for the semantically reliable animate nouns better than the inanimate ones. As predicted, the difference between the red and blue bars in Figure 4 also shows us that participants in the reduplicative agreement condition seem to have learned the class for inanimate nouns better than participants in the other two conditions. Furthermore, the discrepancy between animate and inanimate trials seems largest in the no agreement condition, in line with our prediction that this would yield the lowest category-learning success.

The results of the noun class test were then analyzed using a logistic mixed effects regression model with the lme4 package (Bates et al., 2015) in R (R Core Team, 2018). Condition, animacy and their interaction were set as fixed effects with by-participant random intercepts and by-participant random slopes. Condition was Helmert coded, with no agreement as the reference level; inanimate trials were set as the reference level for animacy using simple treatment coding.

**Noun Class Model:**  $\text{correctness} \sim \text{condition} * \text{animacy} + (1 + \text{animacy} | \text{id})$

The results of this model are summarized in Table 3. The significant intercept ( $b=1.62$ ,  $SE = 0.26$ ,  $z\text{-value} = 6.27$ ,  $p<0.001$ ) confirms that the mean of all condition means for an inanimate trial lies well above the chance level. As expected, participants on average performed significantly better in animate trials than in inanimate ones ( $b = 0.97$ ,  $SE = 0.27$ ,  $z\text{-value} = 3.53$ ,

$p < 0.001$ ). Additionally, a significant negative interaction is shown between the reduplicative agreement condition and animate trials ( $b = -0.73$ ,  $SE = 0.37$ ,  $z\text{-value} = -2.00$ ,  $p = 0.045$ ). This indicates that the mean difference between animate and inanimate trials in reduplicative agreement is smaller than the mean effect of animacy in no agreement and arbitrary agreement combined. No significant difference can be seen for the comparison between the no agreement and arbitrary agreement. However, the negative interaction between reduplicative agreement and animate trials is in line with our primary hypothesis that reduplicative agreement is a particularly helpful type of agreement in learning noun subclasses.

|                                         | Estimate Std | Error  | z-value | p-value     |
|-----------------------------------------|--------------|--------|---------|-------------|
| (Intercept)                             | 1.6210       | 0.2584 | 6.273   | <.001 ***   |
| Animacy Animate                         | 0.9703       | 0.2748 | 3.530   | < 0.001 *** |
| No Agreement vs. Arb. Agreement         | 0.2154       | 0.3646 | 0.591   | 0.555       |
| (No Agr. + Arb. Agr.) vs Red. Agreement | 0.3711       | 0.3665 | 1.013   | 0.311       |
| Arb. Agreement: Animate                 | -0.4618      | 0.3684 | -1.254  | 0.210       |
| Redup. Agreement: Animate               | -0.7330      | 0.3664 | -2.001  | 0.045 *     |

Table 3: Generalized linear mixed model fit by maximum likelihood with results in logits (log likelihood = -1358.0).

#### *Noun Class Learning in Agreement Test*

For interest's sake, we also looked at the correct use of noun class markers on the noun in the agreement test. While these results are not directly relevant to our hypothesis regarding the learning of noun class across different agreement conditions, this test reveals whether participants respond to class any differently when also having to provide the agreement on numerals. The overall performance of participants in this test was slightly higher with an 85% accuracy rate across conditions. Figure 5 below shows that the differences between conditions are reduced here.

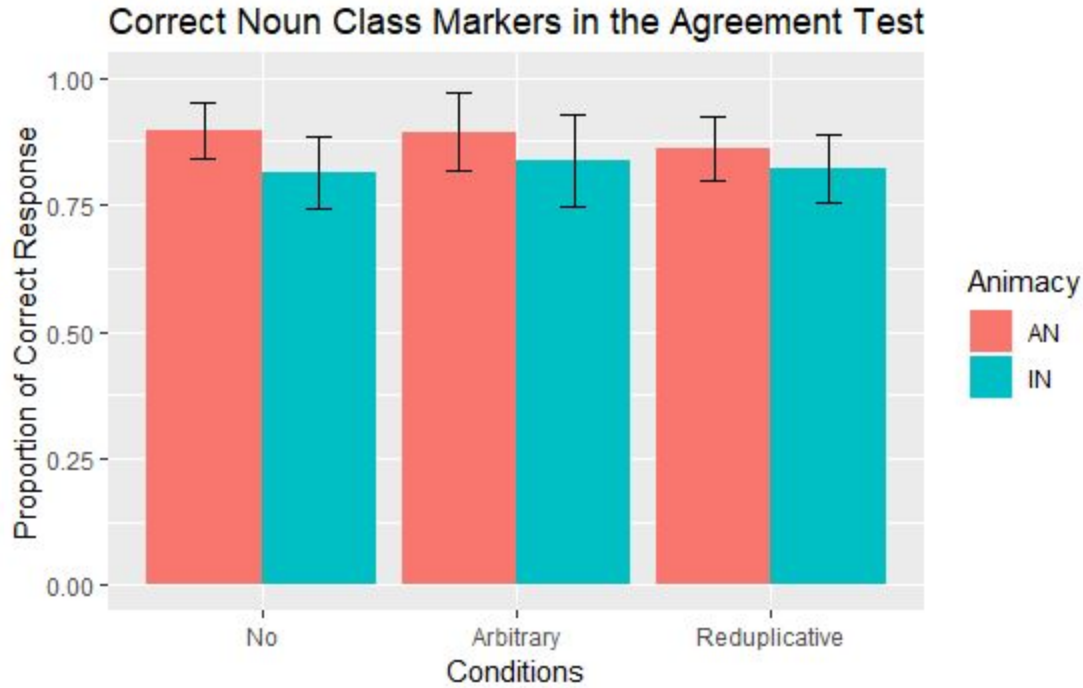


Figure 5: A comparison of participant performance in the three morphological conditions in the agreement test (looking at the use of the markers on the noun). Red bars show the proportion of correct responses for animate trials, blue bars indicate correct responses for inanimate trial. Error bars show a 95% confidence interval for each condition.

The same Noun Class Model (as was used in the noun class test above) was applied to results of the agreement test. As the results below show, only the intercept ( $b = 1.72324$ ,  $SE = 0.25776$ ,  $z\text{-value} = 6.686$ ,  $p < 0.001$ ) and the effect of animate trials remain significant ( $b = 0.95752$ ,  $SE = 0.24443$ ,  $z\text{-value} = 3.917$ ,  $p < 0.001$ ). However, the previously observed effect of the reduplicative condition during animate trials disappears in this test.

|                                         | Estimate Std | Error   | z-value | p-value    |
|-----------------------------------------|--------------|---------|---------|------------|
| (Intercept)                             | 1.72324      | 0.25776 | 6.686   | <.001 ***  |
| Animacy Animate                         | 0.95752      | 0.24443 | 3.917   | < 0.001*** |
| No Agreement vs. Arb. Agreement         | 0.18120      | 0.36545 | 0.496   | 0.620      |
| (No Agr. + Arb. Agr.) vs Red. Agreement | 0.09084      | 0.36337 | 0.250   | 0.803      |
| Arb. Agreement: Animate                 | -0.23207     | 0.33466 | -0.693  | 0.488      |
| Red. Agreement: Animate                 | -0.45955     | 0.32209 | -1.427  | 0.154      |

Table 4: Generalized linear mixed model fit by maximum likelihood with results in logits (log likelihood = -1283.5).

When the testing stage was added as a fixed effect into the model, no significant effect of testing stage was revealed in comparison between noun class test results and agreement test results ( $b=0.08192$ ,  $SE = 0.07178$ ,  $z = 1.141$ ,  $p = 0.254$ ). This suggests that, while the results for noun class learning are different between both testing stages, this difference in performance is not statistically significant.

### *Agreement Learning*

We next turn to the learning of agreement in the two agreement conditions. As can be seen in Figure 6 below, participants performed much better in learning the agreement in the reduplicative agreement condition than with arbitrary agreement.

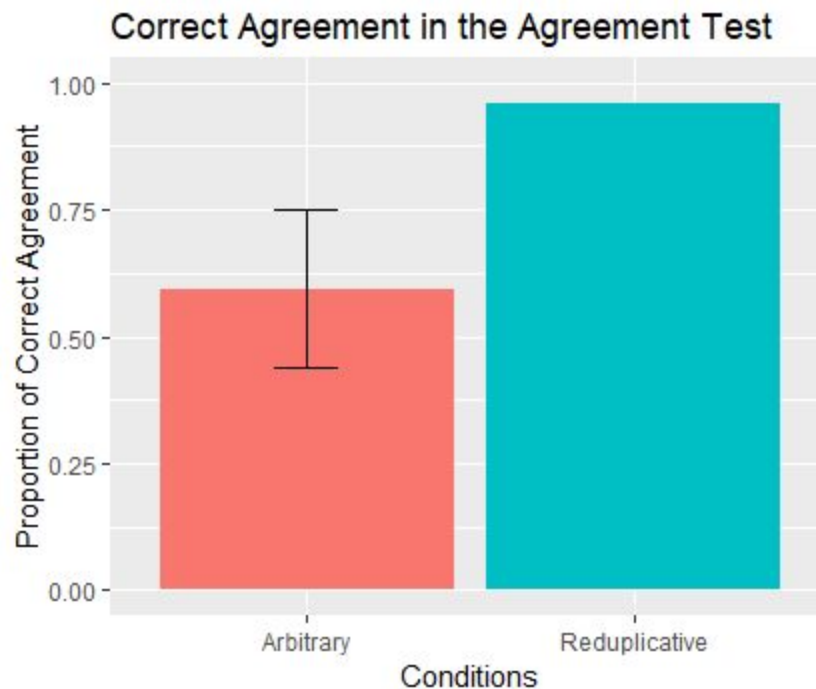


Figure 6: Comparison of agreement learning between arbitrary agreement (red) and reduplicative agreement (blue). The error bar indicates a 95 % confidence interval for responses in the arbitrary condition (no confidence interval was produced for the reduplicative agreement condition due to the near-perfect performance in this condition).

For the analysis of agreement learning, we ran a logistic mixed effects regression model with only condition as a fixed effect comparing reduplicative agreement to arbitrary agreement (the question of agreement was irrelevant for no agreement). Here, reduplicative agreement was set as the reference level and participants were set as random effects.

Agreement Model:  $\text{agreement} \sim \text{condition} + (1 \mid \text{id})$

As expected, participants in the reduplicative agreement condition outperformed participants in the arbitrary agreement condition. The intercept is highly significant, again showing that participants performed well above chance and a high z-value due to the near-perfect performance of participants in the reduplicative condition ( $b = 8.32$ ,  $SE = 0.001$ ,  $z\text{-value} = 7940.88$ ,  $p\text{-value} < 0.001$ ). Additionally, a significant negative effect can be seen for arbitrary agreement ( $b = -7.45$ ,  $SE = 0.79$ ,  $z\text{-value} = -9.45$ ,  $p < 0.001$ ), suggesting that participants performed significantly worse in this condition when compared to the reference level of reduplicative agreement.

|                               | Estimate Std | Error    | z-value  | p-value     |
|-------------------------------|--------------|----------|----------|-------------|
| (Intercept)                   | 8.326475     | 0.001049 | 7940.882 | < 0.001 *** |
| Condition Arbitrary Agreement | -7.450447    | 0.788469 | -9.449   | < 0.001 *** |

Table 4: Model fit for correct agreement with the reduplicative agreement condition as the intercept (log likelihood = -587.2).

To see whether participants in the arbitrary agreement condition did perform above chance in achieving the correct agreement, an intercept-only logit regression model was run using only the data from the arbitrary agreement. This showed that performance in this condition was not above chance as the intercept was not significant ( $b = 0.8209$ ,  $SE = 0.6005$ ,  $z\text{-value} = 1.367$ ,  $p = 0.172$ ).

## *Discussion*

In summary, the results are in line with several of our predictions. Across all conditions, participants performed better in animate trials than in inanimate ones, with the smallest difference observed between trial types in the reduplicative agreement condition. This is reflected in the results of the noun class test, where a significant interaction between animacy and reduplicative agreement was found. Finally, participants learned the agreement between markers on nouns and numerals significantly better with reduplicative agreement than arbitrary agreement. These findings seem to support our hypothesis that reduplicative agreement morphology provides better cues in learning noun subclasses than arbitrary agreement or no agreement, and that reduplicative agreement is easier than arbitrary agreement (in fact, reduplicative agreement seems trivially easy, with 28/30 participants achieving 100% accuracy). However, no significant difference was discovered between arbitrary agreement and no agreement, which does not confirm our overall hypothesis that morphological agreement is better than no agreement altogether.

This may suggest two things: it could indicate a need for adjusting the experimental design or it could indicate a necessary revision of our overall hypothesis. As the above-chance performance across all conditions shows, despite having to learn quite a large language (a total of 30 nouns, two numerals and up to four class markers), participants learned the noun subclasses and morphological agreement surprisingly well. This overall learning success could have weakened

possible effects of morphology in assisting learning as the ease of the task might have enabled memorising of nouns and corresponding markers, rather than encouraging learners to rely on other learning strategies.

In order to adjust the experimental design and increase the difficulty of the experiment in a future study, one could therefore introduce slight alterations to the different experimental stages. One could increase the size of the language by increasing the number of nouns, and/or increasing the number of noun classes, which would also increase the number of markers used. Furthermore, as in Frigo and McDonald (1998), generalization trials could be included to see whether participants learned based on memory or, whether they learned based on semantic and morphological cues. The current discrepancy between animate and inanimate trials across conditions suggests that participants in this experiment did rely on the partially reliable semantic cue in figuring out the noun class system. However, instead of giving exposure of nouns in all possible numerical conditions at the start, and thus provide exposure to all instances of morphological agreement, participants could be exposed to singular and the numeral for two only. In the subsequent testing stages, one could then introduce another numeral word, e.g three. In the agreement conditions, participants would then have to generalise knowledge of both the noun class and the agreement. Such generalisation trials would thus increase the experiment's difficulty and might enable a clearer distinction between the effect of agreement and no agreement conditions.

Secondly, the results might also suggest a necessary revision of the overall hypothesis that morphological agreement is better than no agreement for the learning of noun subclasses. As Table 1 demonstrated in the comparison of cost and benefit of each morphological condition, arbitrary agreement comes with as many costs as no agreement. In other words, while participants have more cues for noun subclass in the arbitrary agreement condition, they also have the cost of learning the agreement between markers. Vice versa, participants in the no agreement condition do not have to struggle with learning morphological agreement, but also have fewer evidence for the categorisation of the nouns. Our results do not show clear support for our hypothesis that one cost/benefit is higher than the other when comparing no agreement to arbitrary agreement. The graphs comparing learning between no agreement and arbitrary agreement seem to suggest a trend in favour of arbitrary agreement, which would suggest that agreement is more important in acquisition. However, the below-chance performance of agreement learning in the arbitrary agreement condition shows that this presents quite a high learning challenge, one that is not there for learners in the no agreement condition.

Considering the above mentioned wealth of evidence highlighting the function of noun class agreement in language processing, these results could indicate that the benefit of agreement morphology primarily lying in its role during communication. Reduplicative agreement may make the acquisition of this communicative function as simple as possible, as it presents learners with a system that is similarly compressible as no agreement in terms of the amount of markers learned. Unlike no agreement though, reduplicative agreement still maximises the

communicative potential of the markers by providing learners and language users with more cues for class. What this might mean is that agreement morphology as such may not be an advantage in the acquisition of noun classes, as in our tests arbitrary agreement morphology shows no clear advantage over no agreement.

This could explain why reduplicative morphology is such a common feature in Bantu languages, as these demand the acquisition of some of the world's most elaborate noun class systems that include up to 23 classes (Demuth 2012). However, it would be noteworthy to investigate why there is no levelling of all agreement morphology to become reduplicative, as opposed to arbitrary. In other words,

Finally, the differences between results of the first and second testing stage might also offer grounds for further research. The main research question aimed to investigate the learning of categories for each noun. For this purpose, results of the first testing stage, where participants had to provide the correct class marker for the noun in isolation, offer the clearest insight into the learning of subclasses. However, the research question also focused on the role of agreement in the learning of noun classes. In the second test, where the conditions' differences in morphological agreement became relevant, differences in category learning became less clear. While it would be premature to base any research assumptions on these results, it would be worthwhile investigating whether future experiments reveal the same pattern between tests on nouns in isolation and tests on nouns in agreement with modifiers.

## *Conclusions*

This study set out to test the following predictions - whether agreement morphology is more helpful than no agreement in the learning of noun classes and, specifically, whether reduplicative morphology is a particularly helpful type of agreement during acquisition. Our results offer support for the latter, i.e. that reduplicative morphology provides better cues for learning noun class than arbitrary agreement or no agreement. However, the results do not confirm the benefit of morphological agreement in noun class learning as such; no clear differences were found between conditions of arbitrary agreement and no agreement. Additionally, stark differences were found between the learning of morphological agreement in the reduplicative agreement condition and the learning of arbitrary agreement. This suggests that reduplicative agreement offers learners one of the simplest mechanisms of acquiring noun class, while arbitrary agreement and no agreement present learners with a similar amount of learning costs. Consequently, no evidence was found for the superiority of agreement morphology as such as a learning mechanism in the acquisition of noun classes.

## *Acknowledgements*

Many thanks to Dr. Jennifer Culbertson and Prof. Kenny Smith for the advice and supervision. Stimuli (words and images) were used from a previous experiment by these advisors and the majority of the PsychoPy script was provided by Jennifer Culbertson, with minor adjustments done by the author.

Recruitment, the running of experiments, transcription and analysis was done by the author. Examples from German and Oshikwanyama were also provided by the author's own knowledge of the languages.

## References

- Anon, I., & Ramscar, M. (2012). Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. *Cognition*, 122(3), 292-305.
- Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and speech*, 47(1), 31-56.
- Bates, D. (2010). lme4: Mixed-effects modeling with R <<http://lme4.r-forge.r-project.org/book>>.
- Braine, M. D. (1987). What is learned in acquiring word classes: A step toward an acquisition theory. *Mechanisms of language acquisition*, 65, 87.
- Caballero, G., & Kapatsinski, V. (2015). Perceptual functionality of morphological redundancy in Choguita Rarámuri (Tarahumara). *Language, Cognition and Neuroscience*, 30(9), 1134-1143.
- Corbett. (1991). *Gender* (Cambridge textbooks in linguistics). Cambridge University Press.
- Culbertson, J., Gagliardi, A., & Smith, K. (2017). Competition between phonological and semantic cues in noun class learning. *Journal of Memory and Language*, 92, 343-358.
- Culbertson, J., Koulaguina, E., Gonzalez-Gomez, N., Legendre, G., & Nazzi, T. (2016). Developing knowledge of nonadjacent dependencies. *Developmental psychology*, 52(12), 2174.
- Culbertson, J., Jarvinen, H., Haggarty, F., & Smith, K. (2018a). Children's sensitivity to phonological and semantic cues during noun class learning: evidence for a phonological bias.
- Culbertson, J., Smith, K., Jarvinen, H., & Haggarty, F. (2018b). Do children privilege phonological cues in noun class learning?
- Demuth, K., & Weschler, S. (2012). The acquisition of Sesotho nominal agreement. *Morphology*, 22(1), 67-88.
- Demuth, K. (2000). Bantu noun class systems: Loan word and acquisition evidence of semantic productivity. *Classification systems*, 270-292.
- Denny, J. P. (1976). What are noun classifiers good for?. In *Papers from the Regional Meeting. Chicago Ling. Soc. Chicago, Ill.* (No. 12, pp. 122-132).
- Dye, M., Milin, P., Futrell, R., & Ramscar, M. (2016). A functional theory of gender paradigms. *Morphological paradigms and functions. Leiden: Brill*.
- Frigo, L., & McDonald, J. L. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language*, 39(2), 218-245.
- Gagliardi, A., & Lidz, J. (2014). Statistical insensitivity in the acquisition of Tsez noun classes. *Language*, 90(1), 58-89.
- Gahl, S., & Garnsey, S. (2006). Knowledge of grammar includes knowledge of syntactic probabilities. *Language*, 405-410.
- Grinevald, C. (2000). A morphosyntactic typology of classifiers. *Systems of nominal classification*, 4, 50-92.
- Karmiloff-Smith, A. (1981). *A functional approach to child language: A study of determiners and reference* (Vol. 24). Cambridge University Press.

- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87-102.
- Levy, Y. (1983). The acquisition of Hebrew plurals: The case of the missing gender category. *Journal of child language*, 10(1), 107-121.
- Manin, D. (2006). Experiments on predictability of word in context and information rate in natural language. arXiv preprint cs/0612136.
- Mills, A. E. (1985). The acquisition of German. In D. Slobin (Ed.). *The crosslinguistic study of language acquisition* (Vol. 1). Hillsdale, NJ:
- Pérez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language*, 18(3), 571-590.
- Rodina, Y., & Westergaard, M. (2012). A cue-based approach to the acquisition of grammatical gender in Russian. *Journal of child language*, 39(5), 1077-1106.
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Senft, G. (2000). What do we really know about nominal classification systems. *Systems of nominal classification*, 4, 11.
- Tsang, C., Chambers, C. G., & Mozuraitis, M. (2017). Compounds, competition, and incremental word identification in spoken Cantonese. *Language, Cognition and Neuroscience*, 32(1), 69-81.
- Smoczyńska, M. (1985). *The acquisition of Polish*. Lawrence Erlbaum Associates, Inc.
- Taraban, R. (2004). Drawing learners' attention to syntactic context aids gender-like category induction. *Journal of Memory and Language*, 51(2), 202-216.
- Tsang, C., & Chambers, C. G. (2011). Appearances aren't everything: Shape classifiers and referential processing in Cantonese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(5), 1065.