

ARTICLE

Poetics of reduplicative word formation: evidence from a rating and recall experiment

Gerrit Kentner^{1,2*}, Isabelle Franz^{1,2} and Winfried Menninghaus¹

¹Department of Language and Literature, Max Planck Institute for Empirical Aesthetics; ²Department of Linguistics, Goethe University Frankfurt

*Corresponding author: Email: gerrit.kentner@ae.mpg.de

(Received 16 July 2021; Revised 17 December 2021; Accepted 20 December 2021)

Abstract

Reduplicative words like *chiffchaff* or *helter-skelter* are part of ordinary language use yet most often found in substandard registers in which attitudinal and expressive meaning components are iconically foregrounded. In a rating experiment using nonwords that either conform to, or deviate from, conventional reduplicative patterns in German, the present study identified affective meaning dimensions, judgments of familiarity and esthetic evaluations of sound qualities associated with such words. In a subsequent recall test, we examined the respective mnemonic potential of the different types of reduplication. Results suggest that, in the absence of semantic content, reduplicative forms are inherently associated with several affective meaning associations that are generally considered positive. Two types of reduplicative patterns, namely full reduplication and [i-a]-vowel-alternating reduplication, boost these positive effects to a particularly pronounced degree, leading to an increase in perceived euphony, funniness, familiarity, appreciation, and positive belittling (cuteness) and, at the same time, a decrease in arousal. These two types also turn out to be particularly memorable when compared both to other types of reduplication and to non-reduplicative structures. This study demonstrates that reduplicative morphology may in and of itself, that is, irrespective of the phonemic and the semantic content, contribute to the affective meaning and esthetic evaluation of words.

Keywords: reduplication; rhyming; vowel alternation; word formation; euphony

1. Introduction

The discipline of poetics (as first defined in Aristotle's *Poetics* (1961)) is about the selection and combination of plot/content features as well as formal linguistic features that render texts of different genres and also single sentences (proverbs, advertisements, etc.) more affectively engaging, stimulating, esthetically liked/preferred and more memorable than other texts or sentences. The present study is about the potentially distinctive affective and cognitive meaning effects as well as the

esthetic evaluations that are associated with single words of a self-reduplicative morphology, such as *dumdum*, *helter-skelter*, or *chiffchaff*. We experimentally investigated, with nonword stimuli based on a small subset of the German phoneme inventory and with German-speaking participants, the expected effects as a function of the varying selection and combination of a set of phonological features that account for different types of word-internal reduplication.

To start with, repetition is one of the most basic linguistic operations. In this study, we focused on repetition on the level of single words, that is, word-internal reduplication. According to Frege's principle of compositionality, the meaning of complex words can be understood by integrating the meaning of their different morphological parts into a compound meaning. In the case of word-internal reduplication, however, the morphological parts do not differ. Rather, the respective words are formed by combining a stem with either a full or a partial copy of itself. What extra meaning does such reduplicative word formation engender? Are the hypothetical meaning effects of forming reduplicative words identical for all such words, irrespective of their morpho-phonological properties? Or do specific linguistic properties support different meaning effects of reduplicative words?

1.1. The iconicity of reduplication

In this study, we experimentally tested the assumption that word-internal reduplication may in and of itself have 'iconic' effects as defined by Peirce (Atkin, 2013), that is, that it directly informs the meaning of the respective words (Fischer, 2011; Kouwenberg & LaCharité, 2001), compare the notion of 'sound-symbolism' (Elsen, 2017; Nuckolls, 1999).

The most obvious iconic meaning association of reduplication can be traced to the concept of *quantity increase*: more of the same form implies more of the same meaning. However, Fischer (2011) has observed that the iconic semiosis of quantity increase associated with reduplication has different and sometimes opposite semantic effects. Moreover, these effects have been shown to be dependent on the base word that is being reduplicated and on the specific context of use (see, e.g., Dingemanse, 2015, on ideophones). Fischer has also emphasized the connection of reduplication with child language or child-directed language (cf. the repetitive production of syllables pervasive in infantile babbling). From this vantage point, word-internal reduplication may support emotive meanings related to ideas of familiarity, affection, diminution, naiveté, and playfulness or fun.

Moreover, repetition or parallelistic patterning is a hallmark of poetic language use (Fabb, 2015; Görner, 2015; Jakobson, 1960; Menninghaus, Wagner, Hanich, et al., 2017; Menninghaus et al., 2018). Reduplicative word-morphology may be considered as a condensed form of poetic language confined to single words, potentially affecting the esthetic evaluation of the respective words with regard to how euphonious or cacophonous they are perceived.

We expected that, dependent on their specific morpho-phonological constituent parts, different types of reduplicative word-morphology may be associated with distinct iconic meaning effects. We tested this assumption with regard to three types of reduplicative patterns that are conventional in German and that we define in the following section.

1.2. Forms and usage of word-internal reduplication in German and English

In German (as well as in English and other West-Germanic languages), reduplicative words are part of ordinary, everyday language. They are primarily associated with substandard and expressive registers, that is, those in which the attitudinal content of the message is foregrounded and the referential and propositional meaning takes a back seat (Bzdega, 1965; Wiese, 1990). Repetition and reduplication can target a variety of linguistic units (phoneme, syllable, morpheme, word, and even phrases) and hence cannot be accounted for with one type of grammatical analysis only (see Kentner, 2017 for a taxonomy). Recent research has shown that specifically three distinct types of word-internal reduplication are fairly frequently used in everyday German language and can therefore be considered conventionalized. Notably, all three types are attested in English, too, and mostly serve similar functions (Benczes, 2012; Ghomeshi et al., 2004; Green, 2016; Horn, 2018; Minkova, 2002; Thun, 1963).

IDENTICAL CONSTITUENT COMPOUNDING (ICC; Hohenhaus, 2004; Finkbeiner, 2014; Freywald, 2015; for the English equivalent see Ghomeshi et al., 2004; Horn, 2018): ICCs show total, that is, identical reduplication of a word stem and convey a context-dependent prototypicality meaning involving a contrast to a non-prototypical alternative. Thus, the expression ‘rice-rice’ emphasizes that the ‘rice’ referred-to is of the ordinary variety rather than, for instance, ‘Basmati’ rice.

- (1) Nimmst du Reis-Reis oder Basmatireis?
 Take-2SG you rice-rice or basmati-rice
 ‘Do you take rice-rice [standard variety rice] or basmati rice?’

A grammatically distinct but similar type of construction is represented by lexical sequences like *zack-zack*, *hopp-hopp*, *dalli-dalli*, *blah-blah* in which doubling serves to intensify the concept communicated by the simplex item. A few onomatopoeias and ideophones likewise show identical reduplication (*Tamtam* ‘tomtom, fuss’, *Bumbum* ‘bang-bang’, *plem-plem* ‘batty’, *balla-balla* ‘crazy’). The list of fully reduplicative ideophones and onomatopoeias is small and not systematically expandable. In contrast, the morphological process that generates ICCs like (1) is productive, that is, it can be employed systematically to create new forms (Freywald, 2015). For a discussion of the morphological status of these forms vis-à-vis other reduplicative types (see Kentner, 2022; Schindler, 1991).

RHYMING REDUPLICATION (Kentner, 2017): In these forms, a monosyllabic or a disyllabic, trochaic word stem is expanded by a rhyming copy (2). As in most end-rhymes, poems, nursery rhymes, etc., the onset consonants of the stem and the reduplicative copy differ in these words. Usually, the rhyming copy features a labial onset consonant [m,b,p], and coronal [d] is found in loans from English (*superduper*, *oki-doki*).

- (2) Schickimicki (< schick(i)), Kuddelmuddel (<Kuddel), Heckmeck (<?)
trendy type (<*posh*), *hodgepodge* (<*tripe*), *fuss* (<?)

Rhyming reduplications are often used as (mocking) nicknames, as hypocoristics, or as words that imply disorder and confusion (*Heckmeck*, *Kuddelmuddel*). That is,

their iconic meaning can be summarized as affectionate or jocular pejoration. For rhyming reduplication in English (*hurly burly*, *hodge-podge*, *Georgie Porgie*), Benczes (2012) and Rastall (2004) note similar effects, namely, humor, playfulness, depreciation, and diminution.

VOWEL-ALTERNATING REDUPLICATION (Kentner, 2017): In these cases, a monosyllabic or (less often) trochaic stem is preceded or followed by a copy in which the vowel is altered. Like rhyme reduplication, these words convey a seemingly antithetic mix of emotions, namely jocular pejoration or affectionate depreciation. Accordingly, vowel-alternating reduplications can be used as (mocking) nicknames (*Sillesalle* < *Silke* [proper name]; an English equivalent is *Nittynatty* < *Natalie*). They may also express durative or iterative event structure (*ticktack*, *Singsang*; corresponding English examples are *tick-tock*, *singsong*) or disorder, scatter and dispersion (*Krimskrams*, *Wirrarr*; English: *bric-a-brac*, *mingle-mangle*).

- (3) *Hickhack* (<hack), *Mischmasch* (<misch), *pickepacke(voll)* (<packen)
bickering (<peck, chop), *mishmash* (<mix), *chock-full* (< pack)

Vowel-alternating reduplication is formally restricted in that it is exclusively applied to stems with the high front vowel [i] or the low vowel [a]. Furthermore, the sequence of stem and copy is strictly determined by the vowel order [i-a] (Kentner, 2017; Minkova, 2002). Accordingly, stems with [a] show prefixing reduplication (*Hickhack* ~ *hack*), whereas stems with [i] have a suffixed reduplicant (*Mischmasch* ~ *misch*). A strong facilitation for this [i-a]-ordering is demonstrable beyond reduplication, for example, in idiomatic binomial expressions (Cooper & Ross, 1975; Green, 2016; Müller, 1997). It is also attested to across language families (Cabrera, 2017).

The reduplication types (2) and (3) are related to reduplicative phrase formations which feature phrase-internal rhyme (English: *name and shame*, *use it or lose it*; German: *hegen und pflegen*, *mit Sack und Pack*) or vowel alternation (English: *pig in a poke*, *splish and splash*; German: *dies und das*, *fix und foxi*). Typically, such reduplicative phrase compositions render these multi-word expressions more salient and memorable (Benczes, 2019), thereby paving the way for an eventual lexicalization or conventionalization of these phraseologisms. In the same vein, reduplications of the rhyme-type (2) and vowel-alternating reduplications (3) are frequently lexicalized (e.g., as nicknames). Words based on identical constituent compounding (1), on the other hand, are usually not lexicalized (Horn, 2018). Instead, they are created specifically for uses in contexts in which a non-prototypical counterpart to the ICC-prototype is salient.

1.3. The present study

In the following experiment, we investigated, by means of behavioral ratings, affective, cognitive, and esthetic effects that speakers of German associate with different types of reduplicative morphology (Part 1). In a subsequent recall experiment (Part 2 of the study), we tested to what extent these different types of reduplicative (non-)words might also affect access to and retrieval from memory.

2. Affective meaning dimensions, cognitive effects, and esthetic appeal associated with the different types of word-internal reduplications (part 1 of the experiment)

2.1. Methods

The data and the analysis code can be accessed with the following link: https://osf.io/qs82r/?view_only=15dafa6ac42f4640b0efabd76ae7edbf

2.1.1. Stimuli

Stimuli were nonwords ranging from two to four syllables. To disentangle effects of repetition per se and of pattern familiarity, all nonword stimuli were systematically varied regarding their phonemic and prosodic make-up, with the reduplicative structures either conforming to one of the conventional reduplicative patterns (1, 2, 3 above) or systematically deviating from them. In addition, we deployed a non-reduplicative baseline condition.

We generated nonce root syllables with three base phonemes each. The root syllables were consonant-vowel-consonant sequences (CVC), conforming to the structure of simple morphological roots in German (Golston & Wiese, 1998), for example, [jaf], combining one of three sonorant onset phonemes [j], [l], or [m] with one of the vowels [i] or [a] and one of the postvocalic fricatives [f], [s], or [ʃ]. This configuration was chosen as it turned out to be relatively productive for nonword generation; other phonemic combinations, for example, with plosive onset consonants, produced many actual words. Care was taken to exclude roots that are valid lexical items in German, for example, [miʃ] 'to mix' or [las] 'to let'.

Roots were combined in sets of two to yield either (1) fully reduplicating [jaf-jaf] or (2) partly reduplicating root pairs. The latter involves minimal alternation concerning one of the three base phonemes, that is, either reduplication with alternation of the onset [jaf-maf], or vowel-alternating reduplication [jif-jaf], or reduplication with alternation of the postvocalic consonant (assonance) [jaf-jas]. Root pairs in the vowel-alternating stimulus conditions were balanced with respect to the vowel order [i-a] and [a-i]. This way, the set of stimuli covers the three conventional patterns exemplified in (1), (2), and (3), and patterns that are not part of the German grammar yet still reduplicative in nature (assonance with postvocalic consonant alternation, vowel alternation with [a-i] order). Finally, non-reduplicating sets [jaf-liʃ] were included as a baseline condition. With this small set (three possible onset sonorants, two vowels, and three postvocalic fricatives) we have tight control over the phonemic content of the stimulus words. Any effects attributable to the individual phonemes will likely affect all stimulus conditions in a similar fashion.

Root pairs were further modified by adding (or deleting) a root-final reduced vowel schwa [ə]. These operations yielded four prosodic structure types: (a) two monosyllables (CVC-CVC), (b) two trochaic disyllables (CVCə-CVCə), and two combinations of a monosyllabic and a disyllabic stem, that is, (c) CVCə-CVC, and (d) CVC-CVCə.

The stimuli were transcribed according to German orthography, with postvocalic [f] and [s] as double graphs <ff> and <ss> to mark the preceding vowel as short lax vowel. German represents postvocalic [ʃ] as a complex grapheme <sch> which

cannot be doubled morpheme-internally. Simplex stem vowels directly preceding [ʃ] are short and lax¹ in German.

The crossing of the phonemic manipulations concerning the (total, partial, or non-) repetition of base CVC-roots, and the prosodic manipulations concerning the syllabic structure of the resulting stems (i.e., addition or elision of schwa [ə]) yielded $6 \times 4 = 24$ stimulus conditions with varying kinds and degrees of repetition. The final stimulus set comprised 120 stimulus items (see Appendix). The exclusion of roots that correspond to potential lexical items affected more roots with stem vowel [a] (*Masche*, *Masse*, *lasch*, *lass*) than [i] (*miss*, *misch*); as a result, roots with [i] are overrepresented in the final set of nonwords.

2.1.2. Planned contrasts

This set-up of stimulus conditions allows for the testing of a number of planned contrasts to disentangle, in the absence of lexical semantic influences, (a) potentially inherent effects of reduplication per se, (b) effects of the particular prosodic shape of our reduplicative non-words, of their main vowel ([a] vs. [i]), and (c) of their conformity vs. non-conformity to ordinary word formation in German.

Effect of reduplication qua reduplication. The first contrast pits the non-reduplicative pattern in which none of the three base phonemes is repeated (baseline, rightmost column in Table 1) against all stimulus conditions featuring reduplicative patterns. This analysis hence targets the effect of repetition per se, independently of any specific type of reduplication.

Effect of prosodic shape. Irrespective of the type of reduplication, the nonwords have different syllabic structures with two, three, or four syllables and different stress patterns (the four rows in Table 1). The prosodic patterns can be conceived of as either conforming to, or deviating from, two relevant principles of prosodic euphony, viz. prosodic balance (Fodor, 1998; Wiese & Speyer, 2015) and rhythmic alternation (Hayes, 1995; Kentner, 2015; Schlüter, 2009). Stimuli are prosodically balanced if they involve equal-sized constituents (an even number of syllables with the same stress pattern, i.e., either two CVC-monosyllables or two CVCə-trochees, respectively); all other stimulus variants are considered unbalanced according to this criterion. Assuming that each CVC-root features stress on its main vowel, the structures with schwa added to the first CVC root (second and last row) allow for rhythmic alternation of stressed and unstressed syllables (schwa being inherently unstressed). Notably, the patterns in (1), (2), and (3) are always prosodically balanced but not necessarily rhythmically alternating.

Effect of main vowel ([i] vs. [a]). High front vowels [i] and low vowels [a] have repeatedly been found to elicit markedly different iconic effects (Auracher, 2015; Elsen, Németh, & Kovács, 2021; Hoshi et al., 2019; Moos et al., 2014; see Lockwood & Dingemans, 2015 for a review). Among the conditions, in which the main vowel is repeated (full CVC reduplication, onset-alternating reduplication, postvocalic

¹The only exceptions to this generalization appear to be the words *Plüsch* [ply:ʃ] ‘plush’, *Rüsche* [ry:ʃə] ‘frill’, and possibly exceptional spellings of names.

Table 1. Stimulus conditions and example nonwords

Syllabic structure modification	Modification of CVC-reduplicative template					
	Full CVC-reduplication	Onset-alternation	Vowel-alternation		Post-vocalic consonant alternation (assonance)	No reduplication (baseline)
			[i-a]	[a-i]		
No schwa	<u>CVC-CVC</u> <u>jaffjaff</u>	<u>CVC-CVC</u> <u>jaffmaff</u>	<u>C[i]C-C[a]C</u> <u>jiffjaff</u>	<u>C[a]C-C[i]C</u> <u>jaffjiff</u>	<u>CVC-CVC</u> <u>jaffjass</u>	<u>CVC-CVC</u> <u>jaffliss</u>
Schwa on 1st CVC-root	<u>CVCə-CVC</u> <u>jaffejaff</u>	<u>CVCə-CVC</u> <u>jaffemaff</u>	<u>C[i]Cə-C[a]C</u> <u>jiffejaff</u>	<u>C[a]Cə-C[i]C</u> <u>jaffejiff</u>	<u>CVCə-CVC</u> <u>jaffejass</u>	<u>CVCə-CVC</u> <u>jaffeliss</u>
Schwa on 2nd CVC-root	<u>CVC-CVCə</u> <u>jaffjaffe</u>	<u>CVC-CVCə</u> <u>jaffmaffe</u>	<u>C[i]C-C[a]Cə</u> <u>jiffjaffe</u>	<u>C[a]C-C[i]Cə</u> <u>jaffjiffe</u>	<u>CVC-CVCə</u> <u>jaffjasse</u>	<u>CVC-CVCə</u> <u>jafflisse</u>
Schwa on both roots	<u>CVCə-CVCə</u> <u>jaffejaffe</u>	<u>CVCə-CVCə</u> <u>jaffemaffe</u>	<u>C[i]Cə-C[a]Cə</u> <u>jiffejaffe</u>	<u>C[a]Cə-C[i]Cə</u> <u>jaffejiffe</u>	<u>CVCə-CVCə</u> <u>jaffejasse</u>	<u>CVCə-CVCə</u> <u>jaffelisse</u>

consonant-alternation), we therefore compared the stimuli that exclusively feature [i] as main vowel with stimuli that only feature [a] as main vowels (between item effect). We consider conditions in which the main vowel alternates to be neutral with respect to this contrast.

Effect of conventionality of reduplicative pattern. Reduplicative patterns that are systematically used for word formation in German and are in this sense conventional are compared to those which are not. Specifically, the conventional stimulus conditions include (a) ICC (1), (b) onset-alternating (2), and (c) [i-a]-vowel-alternating reduplication (3). Non-conventional reduplication patterns (i.e., those not regularly used in word formation) involve alternation of the postvocalic consonant (Assonance) and vowel-alternating reduplications with the reverse vowel order [a-i].

Effect of vowel vs. consonant alternation. Vowels and consonants crucially differ in terms of their sonorance and hence their acoustic salience, with vowels being more salient than consonants. Among the vowels, the contrast between the cardinal vowels [i] and [a] is the greatest potential contrast, as these vowels differ most strongly in terms of sonorance and intrinsic length (Lisker, 1974; Minkova, 2002). The acoustic differences among the three onset sonorants [j,l,m] and the differences among the postvocalic fricatives [f,s,ʃ] are markedly smaller.

Effect of potential lexicalization. Among the conventionalized patterns, onset-alternating (2) and vowel-alternating reduplications (3) may become part of the lexicon, whereas full reduplication is likely not to be eligible for lexicalization (cf. the usually non-lexicalizable ICCs exemplified in (1)). Moreover, ICC or full reduplication (1) differ from the lexicalizable patterns (2) and (3) in that they have a context-dependent prototypicality reading but no clear valence component (positive vs. negative). Rhyming (2) or vowel-alternating reduplications (3), however, do entail evaluative meaning components that may tentatively be summarized as jocular pejoration.

2.1.3. Rating scales

In order to probe the range of potential iconic and affective meanings of reduplication and the sometimes opposing affective and esthetic effects that may be associated with them, we devised six bipolar rating scales. Essentially, the subjective perception of all colors, shapes, sounds, and so forth can be projected onto the fundamental dimensions of the affectively evaluative space (for this concept, see Norris et al., 2010). Accordingly, individual words of a given language can be assigned, and partly have been assigned (see, e.g., the Berlin Affective Word List, Võ et al., 2009; or the norms for English words by Warriner, Kuperman, & Brysbaert, 2013) distinct valence and arousal levels. To start with, our rating scales therefore included the two most fundamental dimensions of Affective Space, that is, positive vs. negative valence and high vs. low arousal (Norris et al., 2010; Scherer, 2005). In addition, we administered scales that targeted two discrete affectively charged evaluations (funny vs. serious and belittling vs. magnifying), a key dimension of cognitive processing (i.e., familiarity vs. non-familiarity, cf. Reber, Schwarz, & Winkielman, 2004), and

esthetic evaluation (euphonious vs. cacophonous). In the following, we explain the reasons underlying the choice of these rating items.

Valence. Reduplications (esp. rhyme reduplication) are frequently used for nickname formation. Nicknames typically presuppose a high degree of familiarity and sympathy (positive appreciation) with the respective person and express positive associations. At the same time, nicknames can also have depreciative implications. The iconic relation of reduplication with dispersion or scatter (i.e., lack of structure and hierarchy) can in some cases support a depreciative meaning dimension (Fischer, 2011; Regier, 1998). We assessed whether and in what way perceived degrees of appreciation/depreciation correlate with the various types of reduplicative word morphology. The respective rating items were

aufwertend (appreciative)–*abwertend* (depreciative).

Because this bipolar rating scale employed allows to measure intensifications of both a positive and a negative evaluation, it captures the effects of reduplicative word formation on positive vs. negative valence attributions, and we henceforth refer to it for short as a ‘Valence’ scale. Still, it should be noted that a prototypical valence scale differs from the one we employed in that the former uses the items *positive* vs. *negative*, whereas ours focuses on capturing perceived changes (up- vs. downgrading) in valence attributions dependent on the linguistic variables of reduplicative word formation.

Arousal. Surveying reduplicative words, we arrived at the conclusion that reduplicative word structures may in general support effects of both increasing arousal (think of the onomatopoeic word *ding-dong* that mimics the quite literally arousing sound of a door bell) and decreasing it (e.g., *mama*). Under the assumption that these opposite effects may not randomly occur, but be systematically associated with differences in the specific linguistic patterns of word-internal reduplication, we probed which types of reduplicative word formation might drive perceived arousal to higher or lower levels. In this way, we captured the effects of reduplicative word formation on the second fundamental dimension of Affective Space, that is, high vs. low arousal. Arousal was measured by the semantic differential

beruhigend (soothing)–*aufregend* (arousing).

Familiarity. Familiarity is well established as a cognitive variable that is predictive of esthetic evaluation. According to the Cognitive Fluency-hypothesis (Reber, Schwarz, & Winkielman, 2004), higher familiarity supports greater ease of processing, and greater ease of processing tends to support higher esthetic liking. On the other hand, both ambitious artworks and non-standard registers of ordinary language use depart by definition from standard expectations and hence include higher degrees of non-familiarity (Wallot & Menninghaus, 2018). We therefore tested whether reduplication per se or rather specific forms of reduplication reduce or boost perceived familiarity and, by implication, perceived strangeness.

The fact that all stimuli are nonwords will most likely drive ratings toward the latter pole. However, we expected that the different types of reduplicative stimuli we

employed would differ in the degree to which they are perceived as strange or familiar. We assessed this hypothetical effect dimension by the semantic differential

fremd (strange)–*vertraut* (familiar).

Euphony. The perceived euphony of single words or texts has been the topic of several studies on phonological iconicity/sound symbolism. Typically, in these studies, the degree of euphony has been correlated with the phonetic characteristics of individual sounds within the word/text (e.g., Crystal, 1995; Miall, 2001; Priestly, 1994). Here, we are interested in the effect of the morpho-phonological pattern rather than the features of individual sounds. Given the widely acknowledged importance of reduplicative sound patterning in poetry (Menninghaus, Wagner, Wassiliwizky, et al., 2017), proverbs (Menninghaus et al., 2015) and product branding (Argo, Popa & Smith, 2010), we were interested in the overall perceived esthetic sound quality of reduplicative nonwords. We assessed this dimension by the semantic differential

wohlklingend (euphonious)–*übel klingend* (cacophonous).

Funniness vs. seriousness. Reduplicative word forms are often seen as a mimicry of child language and/or as somewhat funny uses of language (Benczes, 2012; Dingemanse & Thompson, 2020). In the same vein, among the 12 words judged as most humorous in the study by Engelthaler and Hills (2018), five feature consonant repetitions (*tit*, *booby*, *nitwit*, *twit*, *bebop*), that is, they can be considered reduplicative in a broader sense. On the other hand, repetitive features like rhyme are known to enhance the perception that aphorisms and proverbs communicate a message which deserves being taken seriously (McGlone & Tofiqbakhsh, 2000; Menninghaus et al., 2015). We therefore examined to what extent the different types of reduplicative words enhance perceived funniness or rather perceived seriousness. We measured this hypothetical dimension by the polar adjectives

ernst (serious)–*späßig* (funny).

Augmentation and diminution. Repetition implies formal augmentation, and reduplication may encode in an iconic way amplification (e.g., Maori *ngaru* ‘wave’–*ngarungaru* ‘large wave’). However, partial or total reduplication of syllables may also be used to express attenuation and (affectionate) diminution (Jurafsky, 1996). The diminutive aspect may be enhanced by the mimicry of child language. We assessed the effect of the various reduplicative patterns on this dimension of affective evaluation by the semantic differential

verkleinernd (belittling)–*vergrößernd* (magnifying).

All scales used were five-point Likert scales the endpoints of which are marked by the polar adjectives introduced above. Such semantic differentials (Osgood, Suci, & Tannenbaum, 1957) have proven useful for elucidating multidimensional qualia profiles of a great variety of phenomena.

2.1.4. Set-up of questionnaire

The stimuli were assigned to four lists, based on their prosodic make-up (corresponding to the four rows in Table 1) and fed into an EFS survey presentation

(Questback, 2017). Participants were randomly assigned to one of the four lists. This way, the syllabic structure, or the prosodic modification, was set up as between-participant variable, whereas the modification of the CVC-template was a within-participant variable.

Each stimulus word was presented on screen and was to be rated on the six bipolar rating scales. For each list, presentation order of the stimuli was randomized for each participant. The presentation was set up as follows:

After reading an introductory slide with greetings and instructions, participants went through three practice trials. Both in these practice trials and the trials that were part of the actual study, they were presented a stimulus word centered in the browser window and asked to read it carefully. Upon pressing a button, the stimulus word was shifted to the top of the browser window, and the six rating scales appeared underneath. The order of the rating scales was identical throughout the experiment (as described above).

Participants were asked to rate the stimulus word on the six scales by clicking one of the five positions on each scale. Only after giving all ratings, participants could press a button to move on to the following stimulus word. After the 30 stimulus words were rated on the six scales, participants were prompted to recall as many stimulus words as possible and write them into a form provided in the browser window.

Finally, participants were asked to judge their knowledge of German on a five-point scale (1: *native-like*, 2–3: *fluent*, 4–5: *basic*) and were invited to participate in a raffle to win one of 25 book vouchers (worth 10 Euro each).

2.1.5. Participants

Participants were recruited from the participant pool of the Max Planck Institute for Empirical Aesthetics, Frankfurt (Germany). In addition, the questionnaire was announced on social media platforms of this Institute and on notice boards at Goethe University Frankfurt, with a link provided to access the survey. All in all, 140 participants accessed the online questionnaire. Of these, 88 participants rated the full set, 95 participants at least half of the items, and 119 participants rated at least one stimulus word. All participants who disclosed the respective self-information reported to be native or at least fluent speakers of German.

2.1.6. Data analysis and factor coding

The data of each of the six rating scales were evaluated using Bayesian generalized mixed models with cumulative link function (log link) for ordinal data. The models were implemented in R (R core team, 2020) using the *bmrs* package (Bürkner & Vuorre, 2019); all models assume uninformative priors. The following hierarchically structured factors were each set up as orthogonal sum contrasts according to the design depicted in Table 1, corresponding to the planned contrasts described in Section 2.1.2:

Within-participant effects

1. REDUPLICATION: reduplicative words (+.5) vs. non-reduplicative baseline (−.5).
2. CONVENTIONAL REDUPLICATION: conventional reduplication pattern (+.5: full reduplication, onset-alternation and [i-a]-vowel-alternating reduplication)

- vs. non-conventional or irregular reduplications (−.5: assonance, [a-i]-vowel-alternating reduplication); non-reduplicative stimuli set to 0.
3. LEXICAL REDUPLICATION: lexicalizable onset alternating and [i-a]-vowel-alternating reduplication (+.5) vs. non-lexicalizable full reduplication (−.5); non-conventional reduplication (with [a-i]-vowel alternation or postvocalic consonant alternation) and non-reduplicative stimuli were set to 0.
 4. VOWEL VS. CONSONANT CHANGE: alternating reduplication with vowel change (+.5) vs. alternating reduplication with consonant change (−.5); non-reduplicative stimuli and full reduplication set to 0.

The reduplicative structures without vowel change (full reduplication, onset-alternating reduplication, postvocalic consonant reduplication) have either only the high stem vowels [i] or only the low stem vowels [a] (apart from schwa). We included this vowel contrast as a covariate, as it is associated with sound iconicity and might thus affect certain rating scales:

5. STEM VOWEL I VS. A: Items with only [i] stem vowels (+.5) vs. items with only [a] stem vowels (−.5); Items with vowel change (including non-reduplicative items) set to 0 (between item factor).

Between-participant effects. The four prosodic group conditions are conceived of as obeying or violating two prosodic euphony principles, viz. BALANCE and RHYTHMIC ALTERNATION (see above). Prosodically balanced items have an even number of syllables (two or four syllables: e.g., *jafflich* or *jischemaffe*), unbalanced ones have three syllables with schwa either on the first or the second CVC-root (e.g., *jaffelisch*, *jischmaffe*). The factor Rhythmic Alternation distinguishes patterns in which stressed syllables alternate with schwa syllables (CVC_a-CVC, CVC_a-CVC_a) and those in which two stressed syllables are adjacent (CVC-CVC, CVC-CVC_a).

6. BALANCE: prosodically balanced (coded as +.5) vs. unbalanced (−.5) conditions (between-participants factor).
7. RHYTHMIC ALTERNATION: rhythmically alternating (+.5) vs. non-alternating (−.5) (between-participants factor).

Finally, we included the interaction term BALANCE:REGULARREDUPLICATION. This is motivated by the fact that the reduplicating patterns identified as morphologically regular (full reduplication, onset-alternating, and vowel-alternating reduplication with [i-a] order) require prosodic balance in normal language use, whereas the non-conventional reduplicative patterns are unusual irrespective of their prosodic shape. This difference between the conditions might also affect the ratings.

8. INTERACTION BALANCE:CONVENTIONALREDUPLICATION

Participant and item were entered as crossed random effects in the models, with intercepts and slopes for the fixed effects 1–5 (since the prosodic factors are between-participant factors, they do not feature in the random effect structure).

2.2. Results

The following plots show the mean ratings, broken down by reduplication type (Fig. 1) and prosodic pattern (Fig. 2).

The plots in Figs. 1 and 2 show that the nonwords probed in this survey evoke an interesting mixture of subjectively perceived qualia: they are judged to be at the same time relatively pejorative (depreciative) and arousing, rather strange, as well as cacophonous, funny, and belittling. The spread of the ratings regarding the scales *euphonious–cacophonous*, *serious–funny*, and *familiar–strange* turn out to be substantially larger than the ones for the *soothing–arousing*, *appreciative–depreciative*, and *belittling–magnifying* dimensions, that is, the former three scales appear to be more sensitive/responsive to the differences between the various stimulus conditions than the latter. In general, however, all stimulus conditions have similar profiles in the plots.

We calculated the within-participant correlations between the six rating scales using the *rmcorr* package (Bakdash & Marusich, 2017) in the statistical computing environment R (R Core Team, 2020). To this end, to suit model assumptions, we treat the ordinal Likert-scale data as interval data. The results are depicted in Table 2.

The correlation coefficients show medium-to-large within-participant correlations between the *euphonious–cacophonous* dimension on the one hand, and the *appreciative–depreciative*, *soothing–arousing*, and *familiar–strange* dimensions on the other. Specifically, stimulus items that are rated as sounding more euphonious (or less cacophonous) tend to be rated as more familiar, more soothing, and more appreciative. This is well in line with the hypothesis of familiarity- and hence

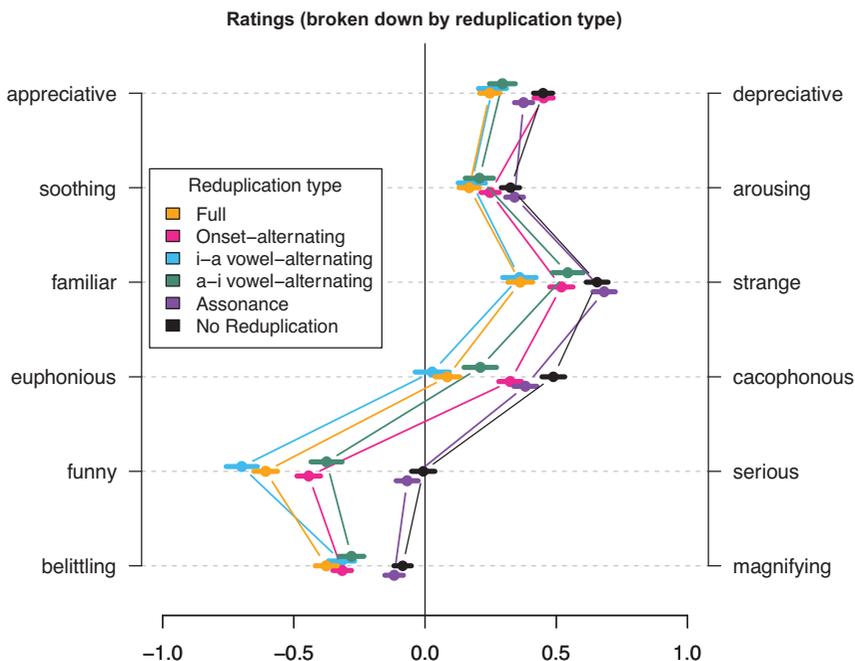


Fig. 1. Ratings by reduplication type. Mean \pm 1 SE. The different prosodic patterns are pooled.

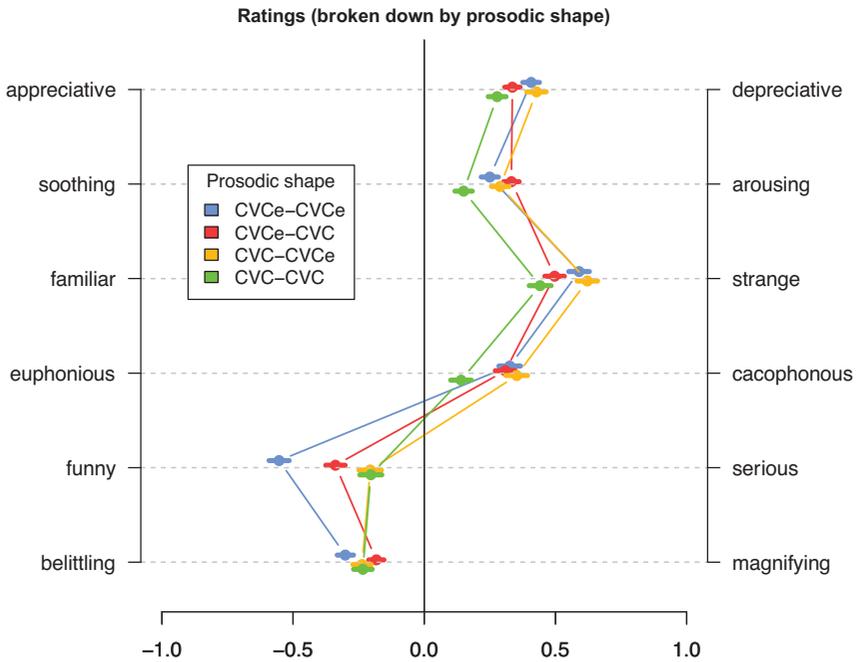


Fig. 2. Ratings by prosodic pattern. Mean \pm 1 SE. The different reduplication types are pooled.

Table 2. Within-participant correlations among the rating dimensions

	Appreciative-depreciative	Soothing-arousing	Familiar-strange	Euphonious-cacophonous	Funny-serious
Soothing-arousing	0.282				
Familiar-strange	0.21	0.22			
Euphonious-cacophonous	0.439	0.312	0.413		
Funny-serious	0.181	0.079	0.236	0.244	
Belittling-magnifying	0.015	-0.009	-0.007	0.052	0.168

cognitive fluency-driven positive esthetic evaluation (Reber, Schwarz, & Winkielman, 2004). There are small to medium-sized correlations between the *funny-serious* scale on the one hand, and all other scales except for the *soothing-arousing* dimension; that is, the perception of funniness regarding these words is related to euphony, familiarity, appreciation, and (affectionate) belittling. Ratings on the *familiar-strange* scale correlate with ratings on the *soothing-arousing* and *appreciative-depreciative* dimensions, with the latter two rating dimensions being correlated as well.

For all these substantial and moderate correlations that hold among the rating scales in general, there are interesting differences regarding the way in which the various stimulus conditions affect the ratings.

In the following, we report for each rating scale the results of a linear mixed model that takes into account the planned contrasts, that is, the effect of (a) reduplication per se, (b) the conventionality or regularity of the reduplicative pattern (conventional reduplication vs. irregular reduplication), (c) the lexicalizability of the productive pattern, (d) and the effects of stem vowel, (e) vowel- vs. consonant-alternation, and (f) prosodic shape.

We first discuss the results for each of the six rating scales. On this basis, we review how the structural differences affect the esthetic and affective evaluation of the stimuli.

In the following plots (see Fig. 3), the eight coefficient estimates (with credible interval that covers 95% of the posterior distribution) of the hierarchical mixed-effects models are depicted for each of the six rating scales. In the following, we specifically discuss the coefficients that considerably deviate from null, that is, when 0 lies outside of 90% of the posterior distribution. The respective coefficients are highlighted in black in the plots.

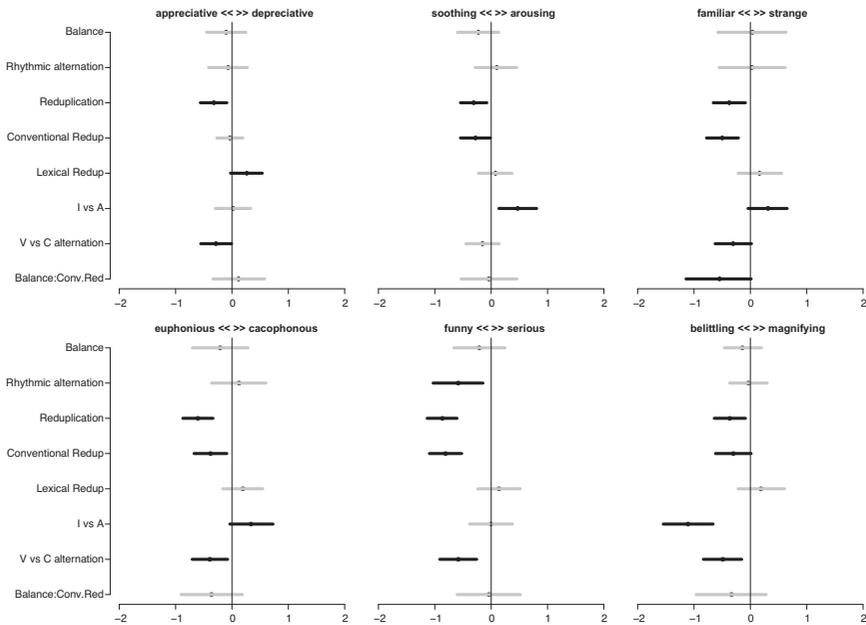


Fig. 3. Coefficient estimates for the ordinal mixed regression models. Coefficient estimates with 95% credible intervals for the Bayesian ordinal mixed-effects models for each of the six rating dimensions, viz. valence (appreciative–depreciative), arousal (soothing–arousing), familiarity (familiar–strange), euphony (euphonious–cacophonous), funniness (funny–serious), and perceived size (belittling–magnifying). Each model has the same eight fixed effects (see subsection on planned contrasts for details): the prosodic shape factors (a) Balance and (b) Rhythmic alternation, and the phonemic factors (c) Reduplication (reduplicative vs. non-reduplicative structure), (d) Conventionality of pattern (conventional vs. non-conventional reduplication), (e) Lexicalizability (lexicalizable vs. non-lexicalizable reduplication), (f) Vowel type (main vowel [i] vs. [a]), (g) Vowel vs. consonant alternation, and (h) the interaction term BALANCE:CONVENTIONALITY of pattern. Coefficients are highlighted (black) if 0 lies outside of the 90% credible interval of the posterior distribution.

2.2.1. Appreciative–depreciative

Reduplicative words are felt to be more appreciative than non-reduplicative ones. However, among the conventional reduplications, the lexicalizable ones (onset-alternating and [i-a]-vowel-alternating reduplication) are rated to be rather depreciative when compared to non-lexicalizable full reduplications. This pejorative effect is mostly due to onset-alternating reduplications, as participants deem vowel-changing reduplications to be significantly more appreciative than consonant-changing ones.

2.2.2. Soothing–arousing

The different stimulus conditions also affect the *soothing–arousing* dimension: Reduplication is felt as rather soothing when compared to non-reduplicative stimuli. This holds especially for the conventional, morphologically regular patterns (full reduplication, onset-alternating reduplication, reduplication with [i-a] alternation). Participants appraise reduplicative patterns that only feature [i] vowels as less soothing or more arousing than the patterns involving [a] as stem vowels.

2.2.3. Familiar–strange

Whereas the nonword stimuli used in this study were in general perceived as rather strange, the reduplicative words have increased familiarity-ratings compared to non-reduplicative structures. Among the reduplications, the conventional patterns are rated as more familiar than the non-conventional ones. Reduplications with two high stem vowels [i] are felt to be stranger or less familiar than patterns with two low stem vowels [a]. Reduplications with vowel change sound more familiar than those with consonant change. The `BALANCE:CONVENTIONALREDUPLICATION` interaction reflects the fact that the familiarity of the conventional or regular patterns specifically increases when they are prosodically balanced (i.e., either two monosyllabic CVC-CVC or two trochees CVCə-CVCə).

2.2.4. Euphonious–cacophonous

The model shows that reduplication per se is felt as relatively euphonious when compared to the non-reduplicative baseline structures. Among the reduplicative structures, the conventional reduplications (full reduplication, onset-alternating, and [i-a]-vowel-alternating reduplication pooled) are valued as more euphonious compared to the non-conventional ones ([a-i]-vowel-alternating, assonance pooled). Reduplicative patterns that only feature [i] stem vowels are perceived as less euphonious than patterns involving two [a] vowels. Among the reduplications with phonemic alternation, vowel change is rated as more euphonious than consonant change. The prosodic factors (balance and rhythmic alternation) do not systematically affect the ratings.

2.2.5. Funny–serious

The structural differences of the stimulus words clearly affect the rating dimension *funny–serious*: The rhythmically alternating conditions are considered funnier than the non-alternating ones. Reduplication per se enhances funniness, and conventional reduplication patterns are felt as funnier than non-conventional patterns. The hypothetically more salient vowel change patterns ([i-a]-vowel alternation and

[a-i]-vowel alternation pooled) score higher on the funny-scale than the consonant-changing reduplications (onset alternation, assonance).

2.2.6. *Belittling-magnifying*

Compared to non-reduplicative words, reduplications are deemed more (positively) belittling. Among the reduplicating structures, this holds specifically for the subset of items the reduplicative pattern of which was identified as regular and conventional (full reduplication, onset-alternating, and [i-a]-vowel-alternating reduplication). The vowel contrast has a particularly strong effect: stems with [i] are perceived as more belittling than stems with [a]. Furthermore, compared to consonant change, vowel change has a stronger belittling effect.

2.3. *Summary and discussion*

Reduplication per se has an effect on all six rating scales: Reduplicative words are felt to be more euphonic, funnier, more familiar, and – to a lesser degree – more soothing, more appreciative and to elicit a stronger effect of (affectionate) belittling when compared to the non-reduplicative baseline words probed in this questionnaire.

Among the reduplicative structures, the conventional patterns (full reduplication, rhyme, and [i-a]-vowel-alternating reduplication) show a stronger effect on the perceptual qualities than the non-conventional reduplications (assonance and [a-i]-vowel alternation). Specifically, the conventional patterns are deemed more euphonic, funnier, more familiar, more soothing, and more (positively) belittling.

Interestingly, the conventionality of the pattern does not appear to systematically affect the *appreciative-depreciative* dimension. This may be due to the fact that the lexicalizable patterns among the conventional ones (especially the onset-alternating reduplication) are perceived as distinctly more depreciative than the non-lexicalizable, fully reduplicating structures. The difference between these two types effectively cancels out an overall effect of valence concerning the conventional patterns. These counteracting effects are likely associated with the functions of the different structures in German: The lexicalizable patterns (rhyme reduplication and vowel-alternating reduplication, see Examples (2) and (3) in the introduction) are often used for jocular depreciation, for example as (mocking) nicknames, while the fully reduplicating ICCs (see Example (1)) are indifferent with respect to the *appreciative-depreciative* scale. The function of ICCs is merely to express prototypicality or intensification, and these are by itself not systematically linked to either appreciation or depreciation. Rather, something can be both prototypically good or prototypically bad, rendering prototypicality-marking reduplications in the end neither a consistent marker of the one nor of the other.

Reduplications with vowel change ([i-a] and [a-i]-vowel alternations pooled) support a rich set of positive enhancing effects when compared to patterns involving a change of a consonant (reduplication with onset-alternation and assonance pooled): They are perceived as more euphonic and funnier, elicit more appreciative responses and are felt to have a more (positively) belittling effect.

Considering reduplications in which the vowels do not alternate (full reduplication, onset-alternation, assonance), stimuli with two [i]s, as compared to stimuli with two [a]s, elicit a strong belittling effect, evoke a higher degree of arousal and sound

distinctly less euphonious. This observation fits well with sound-iconic effects associated with the distinction between low back vowels vs. high front vowels (Auracher, 2015; Dingemanse, 2015; Elsen, Németh, & Kovács, 2021; Hoshi et al., 2019; Shinohara & Kawahara, 2010).

The prosodic manipulations (insertion or omission of schwa [ə]) affected the ratings less strongly compared to the segmental/phonemic manipulations. This weaker effect may be due to the lower statistical power associated with these between-subject factors. Still, the factor RHYTHMIC ALTERNATION does impinge on the ratings for the *funny* vs. *serious* dimension: rhythmically alternating structures are clearly deemed funnier than non-alternating structures. PROSODIC BALANCE alone does not significantly affect the ratings. However, the interaction BALANCE:CONVENTIONAL-REDUPLICATION has an effect on the *familiar-strange* dimension: compared to balanced conventional reduplications (full, onset-alternating, [i-a]-vowel-alternating), unbalanced ones received lower familiarity ratings. This effect is explicable with recourse to the grammatical requirement for the conventional reduplications to exhibit balanced prosodic structure. As one would expect, patterns that are conventionalized in a language should also be more familiar.

3. Recall as a proxy for the memorability of reduplicative structures (part 2 of the experiment)

In the second part of the questionnaire, participants were asked to recall as many of the words they read in the rating study as possible.

All in all, 87 participants provided 582 responses. The mean number of responses was 6.7, ranging from minimally 1 to maximally 20 (median 5, inter-quartile range 3–8.5).

3.1. Data analysis

Responses were normalized for the case, and we determined for each response whether it was a true or a false recall. Regarding the false recalls, we determined whether the response structurally conformed to one of the stimulus conditions of the experiment (see Table 1) or not. We were lenient regarding the orthographic representation of the response, that is, we considered a response as correct if the postvocalic consonant was not written with a double consonant even though it was in the stimulus. However, responses involving consonants or vowels that were not part of the set of stimuli were considered false recalls.

Eight responses that did not observe the CVC(ə)-CVC(ə) structure of the stimulus patterns were marked as invalid and not further analyzed.

3.2. Results

3.2.1. Prosodic structure of responses

All valid responses conformed to the prosodic structure of the stimulus set the participants were presented with (the prosodic structure of stimuli was a between-participant factor and hence did not vary by participant). To determine whether (1) the number of responses and (2) the recall success (number of correct recalls) was dependent on the prosodic structure of the stimuli (Table 3), we applied Chi-square

Table 3. Distribution of recalled items by prosodic structure

	Prosodic pattern			
	CVC _a -CVC _a	CVC _a -CVC	CVC-CVC _a	CVC-CVC
<i>N</i> participants	23	24	18	22
Correct recalls	44	60	48	53
<i>N</i> responses	143	170	118	143

tests. The expected frequency of responses was adjusted to the number of responding participants per prosodic group. Both Chi-square tests yielded non-significant results (total number of responses: $\chi^2 = 1.4$, $df = 3$, p -value = 0.71; number of correctly recalled items $\chi^2 = 2.89$, $df = 3$, p -value = 0.41). Hence the distribution of responses is not likely to be systematically related to the prosodic structure of the items.

3.2.2. Patterns of reduplication

The following Table 4 shows the distribution of valid recalled items broken down by the types of reduplication (the different prosodic structures are pooled).

For the principal data analysis, recalled items that do not conform to any of the stimulus patterns (see 'Other' in Table 4) were disregarded. We will return to a qualitative analysis of these items below.

Apart from the specific stimulus conditions, the number of responses and the recall success are likely to be affected by (a) the presentation ratio (which is lower for the two vowel-alternating conditions compared to the other conditions) and (b) the inherent complexity of the stimuli (with a higher confusion potential for more complex stimuli). We conceive of complexity as a simple function of the number of different phonemes (not counting schwa [ə]), with full reduplication being less complex (three different phonemes) than partial reduplication (four different phonemes) and no reduplication (six different phonemes). We therefore report different Chi-square tests in which the expected frequencies of responses and recalls are adjusted for these factors. In the case of the complexity adjustment, the baseline probability for the expected frequencies is assumed to be proportional to the inverse of the complexity of each pattern (1/3 for full reduplication, 1/4 for partial reduplications, and 1/6 for no reduplication).

Adjusted for both presentation ratio ($\chi^2 = 70.39$, $df = 5$, $p < 0.001$) and stimulus complexity ($\chi^2 = 93.64$, $df = 5$, $p < 0.001$), the χ^2 tests lead us to reject the null hypothesis according to which the distribution of responses is independent of the

Table 4. Distribution of recalled items by reduplicative pattern

	Reduplication Pattern						
	Full redupli- cation	Onset alternating	[i-a]- vowel- alternating	[a-i]- vowel- alternating	Assonance	No redupli- cation	Other
Presentation ratio	0.2	0.2	0.1	0.1	0.2	0.2	0
Correct recalls	78	26	41	18	23	19	-
<i>N</i> responses	121	78	77	26	42	95	135

phonemic pattern. The number of correctly recalled items is not independent of the stimulus patterns either (adjusted for presentation ratio: $\chi^2 = 79.39, df = 5, p < 0.001$; for complexity: $\chi^2 = 38.35, df = 5, p < 0.001$).

Results suggest that, beyond presentation ratio and complexity, the sound patterns of the various stimulus conditions did affect the number of responses and recall success. The plots in Fig. 4 show the Chi-square residuals for the six reduplicative patterns. Positive residuals for any given stimulus pattern suggest that the pattern is

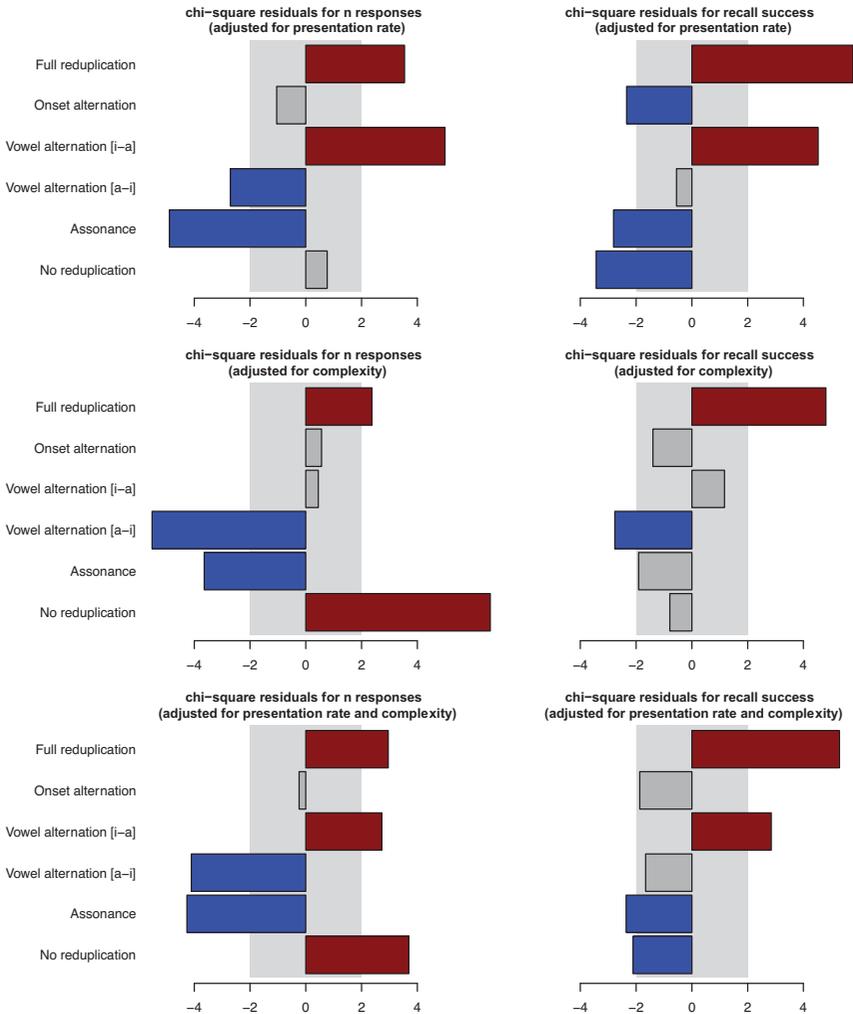


Fig. 4. Chi-square residuals for recalled items. Chi-square residuals for total number of recalled items (left column) and number of correctly recalled items (right column), with expected frequencies adjusted to presentation ratio (upper row), and to pattern complexity (middle row). The bottom row represents residuals adjusted for both, that is, the expected frequency is considered to be commensurate with the mean of the adjustments for presentation rate and complexity. Residuals contributing significantly to the Chi-square distribution, that is, those exceeding $|2|$ are colored blue when negative (indicating underrepresentation), and red when positive (indicating overrepresentation).

overrepresented (i.e., relatively more observed recalls than would be expected when taking into account the ratio of presentation in the rating experiment or the complexity of the pattern), and negative residuals suggest underrepresentation. Chi-square residuals $> |2|$ are considered significant factors for the deviance from homogeneous distribution.

The residuals show that Full Reduplication is overrepresented in the total number of responses as well as in the number of correctly recalled items. Also, vowel-alternating reduplication with the [i-a] ordering is clearly overrepresented when expected frequencies are adjusted to the lower presentation ratio of these items. The assonance pattern, on the other hand, is clearly underrepresented in the responses, and, similarly, recall success is lower than would be expected by chance. Likewise, vowel-alternating reduplications with [a-i] order were recalled by participants less often than expected under the null hypothesis.

The Chi-square residuals for the non-reduplicative patterns and onset-alternating reduplication are less consistent. Non-reduplicative responses are overrepresented when expected frequencies are adjusted for pattern complexity. On the other hand, recall success is clearly diminished for this pattern when the expected frequencies are adjusted for presentation ratio.

In the case of onset-alternating reduplications, recall success is diminished, but only significantly so under the adjustment for presentation ratio.

3.2.3. Effect of stem vowel ([i] vs. [a]) in patterns without vowel alternation

The reduplicative structures without vowel change (full reduplication, onset-alternating reduplication, reduplication with assonance, see Table 5) have either only the high stem vowels [i] or only the low stem vowels [a]. Items with stem vowel [i] were shown more than twice as often as items with stem vowel [a] in the stimulus set. To ascertain whether the vowel affects the distribution of responses and recall success, we applied Chi-square tests in which the expected frequencies are adjusted for the presentation ratio of [i] vs. [a] stems.

For the total number of responses, the significant Chi-square value ($\chi^2 = 53.34$, $df = 1$, $p < 0.001$) suggests that the distribution of responses is not independent of the vowels. A high positive residual (6.2) for stems with [a], and a negative residual (-3.86) for stems with [i] confirm that the former overall facilitated recall as compared to the latter when the unbalanced presentation ratio is taken into account. Among the correct recalls, however, no significant bias for stems with [a] was observed ($\chi^2 = 1.16$, $df = 1$, $p = 0.28$), that is, the number of these recalls can be considered commensurate with the presentation ratio. The striking facilitation for stems with [a] is therefore attributable to a very high percentage of false positive recalls (73%), compared to 51% in case of stems with [i].

Table 5. Distribution of recalled items lacking vowel alternation

Examples	C[a]C(ə)-C[a]C(ə)	C[i]C(ə)-C[i]C(ə)
	<i>lafflaff, maffelaff, jaschjaff</i>	<i>lissliff, miffeliff, jissejisse</i>
Presentation ratio	0.28	0.72
Correct recalls	41	84
<i>N</i> responses	118	122

3.2.4. Other responses

A sizeable number of the items recalled ($n = 143$) did not conform to any of the phonemic patterns of the stimulus categories. The great majority ($n = 135$) of these false recalls still shows the CVC(ə)-CVC(ə) structure yet deviates from the stimuli in that two of the three base phonemes alternate between the two CVC(ə)-roots, whereas, in the original stimuli, either no, only one, or all three phonemes alternate (see Table 6). Seventy-seven of these 135 recalls (57%) involve alternation of both consonants, for example, *mafflass* or *jissemissfe*. The other 58 (43%) feature vowel-alternation plus alternation of one of the two consonants.

Most strikingly, responses with [i-a] vowel order were noted down nearly twice as often ($n = 38$) as responses with [a-i] order ($n = 20$). This response pattern corroborates the facilitation effect for the [i-a] vowel-order that is found and established in the reduplicative pattern with vowel alternation (*Hickhack*, *Singsang*, etc., see Example (3) in the introduction).

3.3. Discussion

Among the nonwords with reduplicative morphology, full reduplication and [i-a]-vowel-alternating reduplication are clearly better stored in, and retrieved from, memory when compared to reduplication involving assonance and [a-i]-vowel alternation. This pattern is mirrored in recall success, which is higher for full reduplication and [i-a]-vowel alternation than for the other conditions. The facilitation of the [i-a]-vowel order is also observable in recalls that, in contrast to the stimulus set, entail two instead of only one phonemic alternation. The preference for the [i-a]-vowel order over [a-i] order is in line with previous findings on the issue (Cabrera, 2017; Cooper & Ross, 1975; Green, 2016; Kentner, 2017; Minkova, 2002; Müller, 1997).²

Moreover, in spite of the lower presentation rate of [a]-stems, we obtained a greater number of recalls with two [a]s when compared to recalls with two [i]s. However, this did not translate into higher numbers of correct recalls, suggesting that

Table 6. Distribution of recalled items involving two alternations

	[i-a]-vowel alternation	[a-i]-vowel alternation	No vowel change
Onset-alternating	$n = 15$, e.g., <i>miffflaff</i>	$n = 9$, e.g., <i>jaffemiff</i>	$n = 77$, e.g.,
Alternation of post-vocalic consonant	$n = 23$, e.g., <i>lischlaffe</i>	$n = 11$, e.g., <i>laffelisse</i>	<i>maffelasse</i>

²One might assume that the [i-a] order could be more frequent in general, leading to the facilitation of this pattern. However, we have no reason to assume a general preference for [i-a] beyond the realm of reduplication and some frozen binomials (e.g., *dies und das* ‘this and that’). For example, the i-derivation/truncation systematically and productively turns German words with [a] as stem vowel into words with the inverse [a-i] vowel order (*Abi*, *Ami*, *Hanni*, *Nanni*, *Mami*, *Papi*, ...). Because the i-derivation/truncation is much more productive and frequent than reduplication, it will easily outweigh the reduplicative [i-a] bias. Therefore, we see no reason to assume such a frequency effect due to the general frequencies in the German lexicon, especially because we are dealing with nonwords.

a facilitation of structures with [a]-vowels do not necessarily enhance correct recalls from memory.

In sum, it appears that full reduplication and [i-a]-vowel alternation have a distinct mnemonic potential.

4. General discussion

In this project, we studied the affective meaning dimensions and the esthetic evaluation associated with various patterns of reduplicative (e.g., *lisslisse*, *miffmaff*, *jischelische*) and non-reduplicative (e.g., *jaffeliss*) nonce words, and explored their respective mnemonic potential.

The results of the rating study show that, overall, the reduplicative nonwords are perceived as being relatively depreciative, arousing, strange, cacophonous, funny, and belittling. At the same time, when compared to non-reduplicative stimuli, reduplicative structure, in general, has a rather positive and hypocoristic effect: it increases the appreciation and the perception of familiarity, euphony, funniness, and (positive) belittling, and it appears to lower the arousing potential of the stimuli. The latter is higher the more depreciative the reduplicative word form is perceived, in line with findings in emotion psychology that negative emotional responses tend to be associated with higher arousal than positive emotions (Baumeister et al., 2001).

However, different kinds of reduplication have markedly different effects on the rating scales. Two kinds of reduplicative nonwords that conform to the phonemic patterns of reduplication found in substandard registers of German (full reduplication, [i-a]-vowel-alternating reduplication) stand out by virtue of being associated with particularly high positive ratings, that is, they are perceived as being more euphonic, funnier, more familiar, more soothing, and more (positively) belittling than the other reduplicative structures. This effect is in good accord with the use of these patterns in substandard German as they are most prevalent in familiar and jocular discourse and – in the case of [i-a]-vowel alternation – serve as a means to create mocking nicknames. These very patterns (full reduplication, [i-a]-vowel alternation) are also overrepresented in the recall experiment. Interestingly, these two kinds of reduplication also elicit comparatively low arousal values.

Conversely, one set of reduplicative nonwords, namely the pattern with alternation of the postvocalic consonant (e.g., *jaffjass*), did hardly affect the ratings when compared to the non-reduplicative stimuli. This assonance pattern turned out to negatively affect recall success. At the same time, this pattern elicits on average the highest arousal ratings.

The relatively low arousal values for highly memorable stimuli and, conversely, relatively high arousal values for stimuli that are more difficult to remember is at odds with the observation that stimuli are usually better memorized when they elicit high arousal (Hourihan, Fraundorf, & Benjamin, 2017; Kleinsmith & Kaplan, 1963). We suggest that the relatively good memorability of full reduplication and [i-a] vowel alternating reduplication is largely due to the low phonemic complexity and the conventionality of the patterns, rather than linked to the arousal values they elicit. The soothingness of these stimuli might also be related to the fact that these patterns are particularly reminiscent of child language (this is plausible when assuming that memory of childhood elicits a feeling of comfort).

Apart from the phonemic manipulation of the CVC template, we explored the effects of prosodic structure, systematically manipulating the number and distribution of stressed and unstressed syllables in the stimuli. This kind of manipulation yielded relatively sparse effects. Most notably, the structures with rhythmic alternation of stressed and unstressed syllables (CVC_σ-CVC_σ, CVC_σ-CVC) were rated as being significantly funnier than patterns involving two subsequent stressed syllables (CVC-CVC_σ, CVC-CVC).

4.1. *Toward a recipe for powerful reduplicative word formation in German*

What are the underlying factors driving the esthetic evaluation and the various expressive and affective meaning dimensions these nonwords are associated with? And what renders some of the stimuli more memorable than others? The pattern of responses suggests four main forces that support higher ratings for euphony, funniness, familiarity, soothing potential, appreciation, and (affectively positive) belittling. The critical features appear to be (a) reduplication, (b) a salient vowel contrast, (c) the regularity or conventionality of the pattern in (substandard registers of) the language, and (d) simplicity or low phonological complexity, that is, absence of phonemic alternation.

The matrix in Table 7 compares the six phonemic patterns with these four factors. The tallies in this matrix group full reduplication with [i-a]-vowel-alternating reduplication on the one hand (three of four features each), and assonance with the non-reduplicative structures (one of four features) on the opposite end of the spectrum. Onset-alternating reduplication and [a-i]-vowel alternation are positioned in the middle of this spectrum (two of four features). This grouping of reduplicative conditions is strongly similar to the ratings obtained in experiment 1 (see Fig. 1). As Fig. 1 shows, among the structures probed, full reduplication and [i-a]-vowel-alternating reduplication, on the one hand, are opposed to reduplication with assonance and non-reduplicative patterns on the other; in fact, this opposition largely holds across all six rating scales.

With these ingredients, we have disclosed a recipe, as it were, for creating memorable, euphonic and funny word forms. The basic ingredient is reduplication. The addition of the salient vowel contrast between [i] and [a] leads to an increase in euphony, funniness, and memorability. Alternatively, the absence of alternation altogether (i.e., full reduplication) amplifies the same qualities. Words that abide by patterns that are conventional means for word formation are considered more euphonic and funnier than reduplications that deviate from the familiar patterns. These reduplicative words turn out to be more memorable as well.

To enhance the funny note, this recipe might be enriched by the insertion of schwa syllables between the CVC roots. The use of stems with [a] instead of [i] (in the cases of reduplication without vowel alternation) will increase euphony (without necessarily increasing funniness) but at the same time reduce the arousal and the belittling effect these words evoke. On the other hand, the positive effects of reduplication all but disappear when reduplication is neither accompanied by vowel alternation nor makes use of a pattern that is regular or conventional in the language, as in the case of reduplication with post-vocalic consonant alternation.

Table 7. Feature matrix for reduplication types

	Full CVC-reduplication	Onset-alternation	Vowel-alternation		Post-vocalic consonant alternation (assonance)	No reduplication (baseline)
			[i-a]	[a-i]		
Example	<i>jaff(e)jaff(e)</i>	<i>jaff(e)maff(e)</i>	<i>jiff(e)jaff(e)</i>	<i>jaff(e)jiff(e)</i>	<i>jaff(e)jass(e)</i>	<i>jaff(e)liss(e)</i>
Reduplication	+	+	+	+	+	–
Vowel contrast	–	–	+	+	–	+
Conventionality of pattern	+	+	+	–	–	–
Simplicity	+	–	–	–	–	–
Tally	3	2	3	2	1	1

5. Conclusion

We investigated, with German-speaking participants, affective meaning dimensions, cognitive judgments of familiarity and strangeness as well as the esthetic evaluation associated with various kinds of reduplicative nonwords, and examined their mnemonic potential. Results suggest that, even in the absence of semantic content, reduplicative forms are inherently associated with several affective meaning associations that are generally considered positive. Two types of reduplicative patterns, namely full reduplication and [i-a]-vowel-alternating reduplication, boost these positive effects to a particularly pronounced degree, leading to an increase in perceived euphony, funniness, familiarity, appreciation, and positive belittling (cuteness) and, at the same time, a decrease in arousal. These two types also turn out to be particularly memorable when compared both to other types of reduplication and to non-reduplicative structures.

References

- Argo, J. J., Popa, M., & Smith, M. C. (2010). The sound of brands. *Journal of Marketing* 74(4), 97–109.
- Aristotle. (1961). *Poetics*. New York: Hill & Wang (S. H. Butcher, Trans.).
- Atkin, A. (2013). Peirce's theory of signs. In E. Zalta (ed.), *The Stanford encyclopedia of philosophy*. Stanford: Stanford University. Retrieved from <https://plato.stanford.edu/archives/sum2013/entries/peirce-semiotics/>.
- Auracher, J. (2015). Synaesthetic sound iconicity: Phonosemantic associations between acoustic features of phonemes and emotional behavior. In M. K. Hiraga, William J. Herlofsky, Kazuko Shinohara, & Kimi Akita (eds.), *Iconicity. East meets West*, 93–108. Amsterdam: John Benjamins.
- Bakdash, J. & Marusich, L. (2017). Repeated measures correlation. *Frontiers in Psychology* 8, 456.
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology* 5(4), 323–370.
- Benczes, R. (2012). Just a load of hibber-gibber? Making sense of english rhyming compounds. *Australian Journal of Linguistics* 32(3), 299–326.
- Benczes, R. (2019). *Rhyme over reason: Phonological motivation in English*. Cambridge: Cambridge University Press.
- Bürkner, P. C. & Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial. *Advances in Methods and Practices in Psychological Science* 2(1), 77–101.
- Bzdęga, A. Z. (1965). *Reduplizierte Wortbildung im Deutschen*. Poznań: Poznanskie Towarzystwo Przyjaciół Nauk.
- Cabrera, J. C. M. (2017). Continuity and change: On the iconicity of ablaut reduplication (AR). In A. Zirkner, Matthias Bauer, Olga Fischer, & Christina Ljungberg (eds.), *Dimensions of iconicity*, 63–84. Amsterdam: John Benjamins.
- Cooper, W. E. & Ross, J. R. (1975). World order. In R. E. Grossmann, L. J. San, & T. J. Vance (eds.) *Papers from the parasession on functionalism*, 63–111. Chicago: Chicago Linguistic Society.
- Crystal, D. (1995). Phonaesthetically speaking. *English Today* 11(2), 8–12.
- Dingemanse, M. (2015). Ideophones and reduplication: Depiction, description, and the interpretation of repeated talk in discourse. *Studies in Language* 39(4), 946–970.
- Dingemanse, M. & Thompson, B. (2020). Playful iconicity: Structural markedness underlies the relation between funniness and iconicity. *Language and Cognition* 12(1), 203–224.
- Elsen, H. (2017). The two meanings of sound symbolism. *Open Linguistics* 3(1), 491–499.
- Elsen, H., Németh, R., & Kovács, L. (2021). The sound of size revisited: New insights from a German–Hungarian comparative study on sound symbolism. *Language Sciences* 85, 101360.
- Engelthaler, T. & Hills, T. T. (2018). Humor norms for 4,997 english words. *Behavior Research Methods* 50(3), 1116–1124.
- Fabb, N. (2015). *What is poetry? Language and memory in the poems of the world*. Cambridge: Cambridge University Press.

- Finkbeiner, R. (2014). Identical constituent compounds in German. *Word Structure* 7(2), 182–213.
- Fischer, O. (2011). Cognitive iconic grounding of reduplication in language. In P. Michelucci, O. Fischer, & C. Ljungberg (eds.) *Semblance and signification*, 55–82. Amsterdam: John Benjamins.
- Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research* 27(2), 285–319.
- Freywald, U. (2015). Total reduplication as a productive process in German. *Studies in Language* 39(4), 905–945.
- Ghomeshi, J., Jackendoff, R., Rosen, N., & Russell, K. (2004). Contrastive focus reduplication in English. *Natural Language & Linguistic Theory* 22(2), 307–357.
- Golston, C. & Wiese, R. (1998). The structure of the German root. In W. Kehrein & R. Wiese (eds.), *Phonology and morphology of the Germanic languages*, 165–185. Berlin: Walter de Gruyter.
- Görner, R. (2015). *Ästhetik der Wiederholung*. Göttingen: Wallstein.
- Green, V. (2016). *Native and non-native intuitions on the phonology of binomial locutions*. Doctoral Dissertation. University of Texas at Austin.
- Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.
- Hohenhaus, P. (2004). Identical constituent compounding – A corpus-based study. *Folia Linguistica* 38(3–4), 297–331.
- Horn, L. (2018). The lexical clone: Pragmatics, prototypes, productivity. In R. Finkbeiner & U. Freywald (eds.), *Exact repetition in grammar and discourse*, 233–264. Berlin: Walter de Gruyter.
- Hoshi, H., Kwon, N., Akita, K., & Auracher, J. (2019). Semantic associations dominate over perceptual associations in vowel-size iconicity. *i-Perception* 10(4), 1–31.
- Hourihan, K. L., Fraundorf, S. H., & Benjamin, A. S. (2017). The influences of valence and arousal on judgments of learning and on recall. *Memory & Cognition* 45(1), 121–136.
- Jakobson, R. (1960). Closing statement: Linguistics and poetics. In T. Sebeok (ed.), *Style in language*, 350–377. Cambridge, MA: MIT Press.
- Jurafsky, D. (1996). Universal tendencies in the semantics of the diminutive. *Language* 72(3), 533–578.
- Kentner, G. (2015). Problems of prosodic parallelism: A reply to Wiese and Speyer (2015). *Linguistics* 53(5), 1233–1241.
- Kentner, G. (2017). On the emergence of reduplication in German morphophonology. *Zeitschrift für Sprachwissenschaft* 36(2), 233–277.
- Kentner, G. (2022). Do not repeat: Repetition and reduplication in German revisited. In M. Eitelmann, & D. Haumann (eds.), *Extravagant morphology*, 181–205. Amsterdam: John Benjamins.
- Kleinsmith, L. J. & Kaplan, S. (1963). Paired-associate learning as a function of arousal and interpolated interval. *Journal of Experimental Psychology* 65(2), 190–193.
- Kouwenberg, S., & LaCharité, D. (2001). The iconic interpretations of reduplication: Issues in the study of reduplication in Caribbean Creole languages. *European Journal of English Studies* 5(1), 59–80.
- Lisker, L. (1974). On ‘explaining’ vowel duration variation. *Glossa* 8(2), 233–246.
- Lockwood, G. & Dingemanse, M. (2015). Iconicity in the lab: A review. *Frontiers in Psychology* 6, 1246.
- McGlone, M. S., & Tofiqbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science* 11(5), 424–428.
- Menninghaus, W., Bohrn, I., Knoop, C., Kotz, S., Schlotz, W., & Jacobs, A. (2015). Rhetorical features facilitate prosodic processing while handicapping ease of semantic comprehension. *Cognition* 143, 48–60.
- Menninghaus, W., Wagner, V., Hanich, J., Wassiliwizky, E., Jacobsen, T., & Koelsch, S. (2017). The distancing-embracing model of the enjoyment of negative emotions in art reception. *Behavioral and Brain Sciences* 40, E347.
- Menninghaus, W., Wagner, V., Knoop, C. A., & Scharinger, M. (2018). Poetic speech melody: A crucial link between music and language. *PLoS One* 13(11), e0205980.
- Menninghaus, W., Wagner, V., Wassiliwizky, E., Jacobsen, T., & Knoop, C. A. (2017). The emotional and aesthetic powers of parallelistic diction. *Poetics* 63, 47–59.
- Miall, D. (2001). Sounds of contrast: An empirical approach to phonemic iconicity. *Poetics* 29(1), 55–70.
- Minkova, D. (2002). Ablaut reduplication in English: The criss-crossing of prosody and verbal art. *English Language and Linguistics* 6(1), 133–169.
- Moos, A., Smith, R., Miller, S. R., & Simmons, D. R. (2014). Cross-modal associations in synaesthesia: Vowel colours in the ear of the beholder. *i-Perception* 5(2), 132–142.
- Müller, G. (1997). Beschränkungen für Binomialbildung im Deutschen. *Zeitschrift für Sprachwissenschaft* 16(2), 5–51.

- Norris, C. J., Gollan, J., Berntson, G. G., & Cacioppo, J. T. (2010). The current status of research on the structure of evaluative space. *Biological Psychology* 84(3), 422–436.
- Nuckolls, J. (1999). The case for sound symbolism. *Annual Review of Anthropology* 28(1), 225–252.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Champaign: University of Illinois Press.
- Priestly, T. (1994). On levels of analysis of sound symbolism in poetry, with an application to Russian poetry. In: L. Hinton, J. Nichols, & J. Ohala (eds.) *Sound symbolism*, 237–248. New York: Cambridge University Press.
- Questback (2017). *EFS survey, version summer 2017*. New York: Questback GmbH.
- R Core Team (2020). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>.
- Rastall, P. (2004). Playful English: Kinds of reduplication. *English Today* 20(4), 38–41.
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review* 8(4), 364–382.
- Regier, T. (1998). Reduplication and the arbitrariness of the sign. In M. Gernsbacher, & S. Derry (eds.), *Proceedings of the 20th annual conference of the cognitive science society*, 887–892. Mahwah: Lawrence Erlbaum Associates.
- Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information* 44(4), 695–729.
- Schindler, W. (1991). Reduplizierende Wortbildung im Deutschen. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 44, 597–613.
- Schlüter, J. (2009). *Rhythmic grammar: The influence of rhythm on grammatical variation and change in English*. Berlin: Walter de Gruyter.
- Shinohara, K., & Kawahara, S. (2010). A cross-linguistic study of sound symbolism: The image of size. *Berkeley Linguistics Society* 36(1), 396–410.
- Thun, N. (1963). *Reduplicative words in English*. Doctoral Dissertation, Uppsala University.
- Võ, M. L., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin affective word list reloaded (BAWL-R). *Behavior Research Methods* 41(2), 534–538.
- Wallot, S., & Menninghaus, W. (2018). Ambiguity effects of rhyme and meter. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 44(12), 1947–1954.
- Warriner, A., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods* 45(4), 1191–1207.
- Wiese, R. (1990). Über die Interaktion von Morphologie und Phonologie – Reduplikation im Deutschen. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 43, 603–624.
- Wiese, R., & Speyer, A. (2015). Prosodic parallelism explaining morphophonological variation in German. *Linguistics* 53(3), 525–559.

A. Appendix

Table A1. Stimulus nonwords used in the experiment

Type	CVC-CVC	CVCe-CVC	CVC-CVCe	CVCe-CVCe
a-i redup	jaffjiff	jaffejiff	jaffjiffe	jaffejiffe
a-i redup	laffliff	laffeliff	laffliffe	laffeliffe
a-i redup	lassliss	lasseliss	lasslisse	lasselisse
i-a redup	jischjasch	jischejasch	jischjasche	jischejasche
i-a redup	jissjass	jissejass	jissjasse	jissejasse
i-a redup	miffmaff	miffemaff	miffmaffe	miffemaffe
Assonance	jaffjasch	jaffejasch	jaffjasche	jaffejasche
Assonance	jassjaff	jassejaff	jassjaffe	jassejaffe
Assonance	jiffjisch	jiffejisch	jiffjische	jiffejische
Assonance	jischjiss	jischejiss	jischjisse	jischejisse
Assonance	liffliiss	liffeliss	liffliisse	liffelisse
Assonance	lischliiss	lischeliss	lischliisse	lischelisse
Rhyme	jiffmiff	jiffemiff	jiffmiffe	jiffemiffe
Rhyme	jissliss	jisseliss	jisslisse	jisselisse
Rhyme	laffmaff	laffemaff	laffmaffe	laffemaffe
Rhyme	liffjiff	liffejiff	liffjiffe	liffejiffe
Rhyme	lissjiss	lissejiss	lissjisse	lissejisse
Rhyme	miffliiff	miffeliff	miffliiffe	miffeliffe
No redup	jafflisch	jaffelisch	jafflische	jaffelische
No redup	jassliff	jasseliff	jassliffe	jasseliffe
No redup	jassmiff	jassemiff	jassmiffe	jassemiffe
No redup	jischmaff	jischemaff	jischmaffe	jischemaffe
No redup	lischjaff	lischejaff	lischjaffe	lischejaffe
No redup	miffjass	miffejass	miffjasse	miffejasse

Note. The columns represent the four prosodic structure types.

Cite this article: Kentner, G., Franz, I. & Menninghaus, W. (2022). Poetics of reduplicative word formation: evidence from a rating and recall experiment. *Language and Cognition* 14: 333–361. <https://doi.org/10.1017/langcog.2021.27>