

JOURNAL ARTICLE

The impact of animacy and speech rhythm on the word order of conjuncts in German preschoolers and adults

Isabelle Franz^{1,2}, Gerrit Kentner^{1,2} and Frank Domahs³

¹ Department of Language and Literature, Max-Planck-Institute for Empirical Aesthetics, Frankfurt, DE

² Goethe University, Frankfurt, DE

³ University of Erfurt, DE

Corresponding author: Isabelle Franz (Isabelle.Franz@ae.mpg.de)

In this study, we investigated the impact of two constraints on the linear order of constituents in German preschool children's and adults' speech production: a rhythmic (*LAPSE, militating against sequences of unstressed syllables) and a semantic one (ANIM, requiring animate referents to be named before inanimate ones). Participants were asked to produce coordinated bare noun phrases in response to picture stimuli (e.g., *Delfin und Planet*, 'dolphin and planet') without any predefined word order. Overall, children and adults preferably produced animate items before inanimate ones, confirming findings of Prat-Sala, Shillcock, and Sorace (2000). In the group of preschoolers, the strength of the animacy effect correlated positively with age. Furthermore, the order of the conjuncts was affected by the rhythmic constraint, such that disrhythmic sequences, i.e., stress lapses, were avoided. In both groups, the latter result was significant when the two stimulus pictures did not vary with respect to animacy. In sum, our findings suggest a stronger influence of animacy compared to rhythmic well-formedness on conjunct ordering for German speaking children and adults, in line with findings by McDonald, Bock, and Kelly (1993) who investigated English speaking adults.

Keywords: Rhythm; animacy; picture naming; word order; children; adults

1. Introduction

The transmission of thought into speech involves the retrieval of appropriate lexical items and their ordering according to the rules of syntax. Syntax, however, does not fully determine word order; instead, speakers often have to decide between possible word order variants when formulating their message. Semantic as well as phonological constraints are known to affect such word order decisions in speech production, and they do so to varying degrees. In normal, spontaneous language use, semantic constraints presumably control word order more immediately and to a stronger degree than phonological constraints. This follows from the logical directionality of language production, in which the semantic content of the message governs lexical choice and the assignment of syntactic function; phonology can exert its role and endow the structure with sound only once a syntactic scaffold has been constructed (Levelt, 1989). Nevertheless, phonological influences on word order are on record (Breiss & Hayes, 2020; Büring, 2013; see Anttila, 2016 for a review), but they appear to be limited to sub-clausal environments (Kentner & Franz, 2019).

However, one field in which phonological influences on sentence structure appear to be comparatively strong is child language. For instance, Gerken (1996) showed that English-speaking toddlers are more likely to produce grammatically ill-formed sentences (omitting

obligatory elements) when this leads to rhythmic and prosodic well-formedness.¹ The apparent strength of prosodic/phonological leverage vis-à-vis syntax in child language might be due to a phonological template in language acquisition (Gerken, 1994a) in combination with syntactic uncertainty. Grammatical rules, like the obligatory presence of a determiner in a phrase like *Mary tickled the doll*, are not yet firmly established. This leaves room for syntactically deviant but phonologically unmarked structures (e.g., promoting rhythmical alternation of stressed and unstressed syllables).

Evidence for the power of prosodic and rhythmic constraints in language production of older children is rather limited. For example, Domahs, Blessing, Kauschke, and Domahs (2016) report a systematic influence of prosody on written production. In their study, second graders avoided prosodic violation at the cost of grammatical rules such as omitting an obligatory determiner. Whereas most grammatical rules and phonological structures should be established at this age, additional task demands, as induced by writing, seem to dig out effects of the underlying prosodic structure. However, as opposed to studies on the omission of obligatory elements, there are no studies concerned with the influence of rhythm on the word order in child language.

As to semantic constraints, the evidence in favor of an effect of animacy on word order in child language is relatively robust: Animate referents tend to be produced before inanimate ones. An influence of this animacy constraint (ANIM) was shown, among others, by Prat-Sala et al. (2000) for English and Catalan children and by Drenhaus and Féry (2008) for German preschoolers. While studies comparing the influence of semantic and phonological constraints on word order have been conducted with adult participants (McDonald et al., 1993) or on corpus data (Lohmann, 2013), their impact and interaction during language development is understudied.²

Here, we set out to test the interactive effects of rhythm and animacy on word order, comparing child and adult speech production. To this end, we emulate the study by McDonald et al. (1993) and elicit coordination structures with two bare nouns as conjuncts, the chosen sequence of the nouns serving as the dependent variable. In three experiments, we pit the semantic constraint ANIM (i.e., the preference to order animate referents before inanimate ones) against the prosodic constraint *LAPSE (i.e., the avoidance of structures involving stress lapses). The results reveal a robust effect of animacy for both adults and children. For both age groups, an effect of *LAPSE was also detectable, but only when the animacy of the nouns involved did not vary.

The article is structured as follows: The remainder of Section 1 reviews previous research concerning the effects of rhythm and animacy on word order in child and adult speech production. Sections 2–4 report on three experiments, one with preschoolers, and two with adult participants. Section 5 provides a detailed discussion and concludes the paper.

1.1. The animacy constraint

One of the cross-linguistically most relevant constraints on the order of nouns in spoken and written language is based on animacy: Nouns with animate referents tend to be produced earlier than those with inanimate referents (Saueremann & Höhle, 2018; Lohmann, 2013; Branigan, Pickering, & Tanaka, 2008; Grewe, 2007; Demuth, Machobane, Moloi, & Odato,

¹ In her studies, Gerken mainly addressed prosodic structure, not just rhythm. However, since her items include stress lapses, her series of experiments is highly relevant here.

² To our knowledge, there is only one study that focusses on an interaction of meter and a non-phonological constraint. Investigating the conditions of children's tendency to omit unstressed function morphemes, Boyle and Gerken (1997) showed that not only rhythm, but also the familiarity of the surrounding nouns or verbs is a crucial predictor for young children's omissions. That is, in phrases with highly unfamiliar words, children tend to omit syllabic function morphemes, whereas in sequences with familiar nouns and verbs, children preserve them. The authors found familiarity to be effective independent of the metrical structure of a phrase.

2005; McDonald et al., 1993). Beyond the simple dichotomy (animate versus inanimate), a more gradual continuum of animacy has been proposed (Paczynsky & Kuperberg, 2011). Such graded rankings can be justified by different criteria (Yamamoto, 1999), with the ability of autonomous movement (self-propelledness) being the most important feature for preschool children (Piaget, 1978; Opfer & Gelman, 2011). However, a dichotomous classification seems sufficient for our purpose.

How does animacy influence word order? A first type of account attributes the effect to lexical or conceptual access. Bock and Warren (1985) attributed facilitated access to animate concepts to the fact that they have more semantic features or pathways, arguing that these features accelerate activation of an item ('concept accessibility'). A second type of explanation highlights the standard order of thematic roles in the lexical entry of a verb and its theta grid. According to this account, agent and experiencer tend to occupy earlier positions and, being capable of self-controlled actions and feelings, usually animate referents fill these positions. This approach can be located at the interface between syntax and semantics (Grewe, 2007). However, Dewart (1979) advanced a purely syntactic explanation according to which the tendency to produce animate items first is considered to be due to the fact that they usually serve as subjects. For SVO-languages like German (which has a non-rigid word order³), this syntactic approach could at least partly explain the animacy constraint.

1.1.1. Animacy and child language

Several studies have shown that infants at an age of eighteen months and possibly even younger are capable of distinguishing animate and inanimate referents (Legerstee, Pomerleau, Malcuit, & Feider, 1987; Poulin-Dubois et al., 1996; Legerstee & Markova, 2008). Older children like preschoolers seem to make very reliable animate-inanimate distinctions (Wright, Poulin-Dubois, & Kelley, 2015). In a study by Gelman and Gottfried (1996), children watched video clips with moving animals or moving objects both being carried by a human hand. After watching those movies, participants had to explain why the referents were moving. Most children remembered the hand as an external reason for the movement of artefacts whereas animals were remembered having moved autonomously. These findings suggest that children presuppose that animate referents are capable of self-paced movement whereas inanimate objects are not. With self-paced movement being the most reliable criterion for the attribution of animacy in children (Piaget, 1978; Yamamoto, 1999), at least an implicit ability to distinguish between animate and inanimate referents can be concluded from the study by Gelman and Gottfried (1996; for an overview, see Opfer & Gelman, 2011).

Several findings suggest some effect of animacy in child language similar to the constraint's impact on adult language (Demuth, 2005). Prat-Sala et al. (2000) asked children aged four to eleven to describe dynamic picture scenes showing animate and inanimate objects. Participants were more likely to produce object topicalizations when objects were animate. In another study, German speaking children aged three to six reproduced sentences with three arguments varying in animacy (Drenhaus & Féry, 2008). When making mistakes in word order, children produced significantly more often arguments with animate referents before those with inanimate referents than the other way around. Byrne and Davidson (1985) reported similar results concerning English and Fijian preschoolers' tendency to sequence animate and inanimate items. Subsequent to the memorization of names referring either to animals or to objects, participants were asked to reproduce those names in pairs. Crucially, names with animate referents were produced before those with

³ Even though German has a non-rigid word order, there are typically variants that are less marked and consequently more frequent than others (for example SO sentences as compared to OS sentences).

inanimate referents⁴ (Byrne & Davidson, 1985). A more recent picture description study by Gàmez and Vasilyeva (2015) supports these results. The authors showed that English speaking preschool children are more likely to produce passive sentences when the patient is animate and the agent inanimate (e.g., *The dog was hit by the car.*) as opposed to a combination of two animates—in these cases children preferred active constructions (e.g., *The man hit the dog.*). In summary, these findings demonstrate that the animacy constraint has some impact on word order in children’s (as well as in adults’) speech production.

1.2. Rhythm and *LAPSE

The phonological factor of interest is the propensity for an alternating rhythm in speech (Principle of Rhythmic Alternation, Selkirk, 1984). In German, as in English, the underlying trochaic pattern of prosodic word structure often yields this alternation of stressed and unstressed syllables in spoken language. Ideally, this alternation is not disturbed as in (1).

(1) *Ríta sórted véry mány bóoklets.*

However, in everyday speech, disruptions of this rhythmic structure are inevitable and do occur naturally. Rhythmic deviations may appear as stress clashes or stress lapses (Selkirk, 1984; Hayes, 1995); the former denoting an encounter of two or more stressed syllables, the latter a sequence of two or more unstressed syllables. A prosodic constraint in speech, *CLASH, requires that sequences of stressed syllables should be avoided. Accordingly, another prosodic constraint, *LAPSE, requires that sequences of unstressed syllables should be avoided. One way to visualize the rhythmic structure of phrases or sentences is a metrical grid (Halle & Vergnaud, 1987). In such a grid, the degree of a syllable’s prominence is shown by the number of beats that are accumulated above it, and rhythmic structure is defined by the horizontal distances of equally prominent syllables. Hence, in a rhythmically alternating structure, the number of beats should alternate from one syllable to the next syllable. According to Halle and Vergnaud (1987), the number of beats is determined by principles of the prosodic hierarchy (see below). Since we are only dealing with the trochaic pattern in (1) as well as its violation in structures with *LAPSE and *CLASH, **Figure 1** is limited to two levels of prominence (i.e., the distinction between stressed and unstressed syllables).

x		x		x		x		x		2
x	x	x	x	x	x	x	x	x	x	1
Ri-	ta	so-	ted	ve-	ry	ma-	ny	book-	lets.	
x		x	x							2
x	x	x	x	x						1
Ri-	ta	sorts	book-	lets						
x		x			x					2
x	x	x	x	x	x	x				1
Ri-	ta	so-	ted	the	book-	lets				

Figure 1: Simplified metrical grids for a rhythmic sentence (upper panel), a sentence with stress clash (middle panel), and a sentence with stress lapse (lower panel). The two rhythmic violations *CLASH and *LAPSE are highlighted in bold. Layer 1 represents the default beat for every syllable. Layer 2 represents the beat for every stressed syllable.

⁴ Interestingly, Fijian languages are subject final (Byrne & Davidson, 1985). This weakens the pure syntactic explanation for ANIM as proposed in Dewart (1979).

1.2.1. *LAPSE in the prosodic hierarchy

A stress lapse is not merely considered a rhythmic imperfection concerning the sequential ordering of stressed and unstressed syllables. Rather, in research on prosodic phonology, a stress lapse necessarily involves a deviance from an optimal hierarchical representation.

In prosodic phonology, utterances are divided into prosodic units, which are layered according to their size. Hence, every utterance can be represented as a hierarchic tree structure (see **Figure 2**). The strict layer hypothesis (Selkirk, 1984) postulates that lower units in the prosodic hierarchy should be exhaustively parsed into the next higher unit ('exhaustivity constraint'⁵). Thus, syllables are dominated by feet (F), which in turn are dominated by prosodic words (PW) and phonological phrases (PHP). For our purposes, the interplay of syllables and feet is most relevant. In German, English, and other languages, syllables are grouped into left-headed, maximally binary feet with the left syllable being stressed (e.g., *Brúder*, 'bróther').

Figure 2 illustrates the prosodic hierarchy and its relation to *exhaustivity* and *LAPSE. As a foot should consist of at most two syllables, with one syllable being stressed, the first three syllables in the example *Lane, kiss* and *-es* can each be parsed into a foot. In terms of *exhaustivity*, they form an optimal structure, as they are parsed exhaustively into the next higher domain, i.e., prosodic feet. In contrast, the fourth syllable *the* cannot be parsed into one of the surrounding trochaic feet and must therefore be adjoined directly to the prosodic word, thereby violating the exhaustivity constraint (marked in red). At the same time, this structure engenders a lapse with two adjacent unstressed syllables. Under the assumption that feet are maximally disyllabic and trochaic, stress lapses generally incur a violation of the exhaustivity constraint. **Figure 2** also illustrates that this doesn't hold for stress clash: Although it causes a rhythmic violation, the first two syllables *Lane* and *kiss* are parsed exhaustively, either forming their own foot or at least forming the head of a foot (see Gerken, 1996).

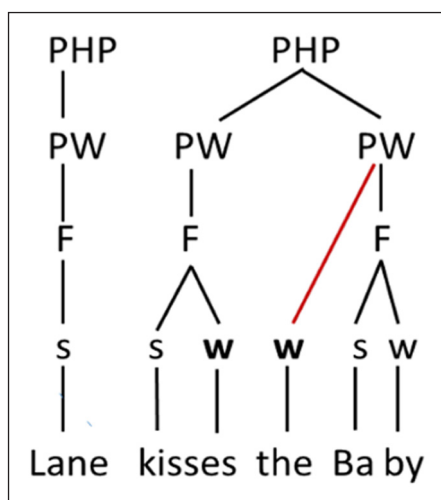


Figure 2: Example for the prosodic hierarchy. The letters s (strong) and w (weak) in the lowest level of the tree label the syllables as being strong/stressed or weak/unstressed. The red line represents non optimal syllable parsing. There is a lapse structure consisting of two weak syllables, *-es* and *the*, highlighted in bold.

⁵ We are aware that *exhaustivity* is considered a violable constraint, as suggested by Wagner (2010), among others. Still, Demuth (2007) showed that children prefer prosodically unmarked structures, i.e., those that best adhere to *strict layering* as envisaged by Selkirk (1984). This involves observing constraints like *exhaustivity*, *layeredness* or *non-recursivity* and is the essence of the *prosodic licensing hypothesis*. We share the view that speakers, especially children, prefer unmarked, i.e., non-recursive, structures when they have a choice. Further, Wagner (2010) stated even for adult language that: "Human beings are not particularly good at processing nested structures and seem to prefer to keep recursive depth at a minimum." (Wagner, 2010, p. 232).

1.2.2. Empirical findings on *LAPSE

The repudiation of structures involving stress lapses is vividly demonstrated by Shih (2014) who shows that forenames are preferably chosen in a way so that they don't lead to lapses in combination with the corresponding surnames. For example, a name like *Denise Fitzgerald* is more likely to be chosen than *Ánnelise Fitzgerald*, as the latter leads to stress lapse with three unstressed syllables in a row (*-ne*, *-lise*, and *-Fitz*) (Shih, 2014, p. 53). There are also studies on rhythmic irregularities in speech processing. For example, Bohn, Knaus, Wiese, and Domahs (2013) showed in an ERP study that the brain is sensitive to stress lapses and clashes when processing German phrases. Further, a corpus study by Breiss and Hayes (2020) suggests that the preference for rhythmic alternation extends to word ordering and sentence formation.

Similarly, Lee and Gibbons (2007) report that English speaking adults avoid lapses when formulating sentences: Participants were more likely to produce the unstressed optional complementizer *that* when it is surrounded by stressed syllables. Thus, the optional complementizer was more often produced in sentences like (2a) than in sentences like (2b).

(2a) *Hénry knéw (that) Lúcy wáshed the dishes.*

(2b) *Hénry knéw (that) Louíse wáshed the dishes.* (Lee & Gibbons, 2007, p. 449)

However, Kentner and Franz (2019) conducted a conceptual replication of Lee and Gibbon's experiment in German and found no such effect. It appears that the power of rhythm is limited: In contrast to English, the difference between introduced and un-introduced German sentences is not just marked by the complementizer (or its absence) but also by the syntactic structure of the sentence such that introduced German sentences are verb-final whereas un-introduced ones are verb-second, as demonstrated in the examples in (3).

(3a) *Sandra glaubt, dass Gisbert Techno hört.* (verb-final CC)
Sandra thinks that Gisbert Techno listens

(3b) *Sandra glaubt, Gisbert hört Techno.* (verb-second CC)
Sandra thinks Gisbert listens Techno
'Sandra thinks (that) Gisbert listens to Techno music.'

Hence, it seems that *LAPSE exerts some power as long as this does not imply any higher-level structural change of the sentence. (For an overview of rhythmic influences on syntactic encoding, see Kentner & Franz, 2019.)

1.2.3. Prosodic constraints and child language

It is well known that infants are capable of recognizing rhythmic structures of their mother tongue (Mehler et al., 1988; Nazzi & Ramus, 2003). Furthermore, prosodic speech processing has been described as an important prerequisite for the acquisition of different linguistic domains. For instance, it has been shown that infants are able to identify lexical or syntactic boundaries based on prosodic cues. In consequence, the ability of infants to segment the speech signal using those cues as bootstraps has been termed *prosodic bootstrapping*, introduced by Pinker (1984) (for toddlers see also Schröder & Höhle, 2011; Eimas, 1996; Morgan & Demuth, 1996; Morgan, 1986; for three-year-old children Männel & Friederici, 2016).

According to the *prosodic licensing hypothesis*, prosodic constraints are also particularly important in toddlers' speech production in that they show a propensity for prosodically well-formed structures (Demuth, 2007). In a seminal series of studies, Gerken (1991, 1994b, 1996) observed that young children reproduce sentences which are more consistent with the prosodic hierarchy and its rules than the actual stimulus. *LAPSE has been found

to be particularly powerful: The determiner *the* was more likely to be omitted in sentences like (4b) than in sentences like (4a), even though the omission leads to ungrammatical sentences (Gerken, 1996).

(4a) *He kícks the píg.*

(4b) *He cácthes the píg.*

Note that the determiner *the* in (4b) can't be parsed into a trochaic foot since the weak position of the trochee is already occupied by the syllabic inflection *-es*. In contrast, in (4a), the inflected verb is monosyllabic and the determiner can prosodically adjoin to it to create a trochaic foot. Accordingly, in (4b) the determiner creates a violation of the exhaustivity constraint (Gerken, 1996). Demuth (2007) replicated these findings and showed that the tendency to omit function words to repair prosodic violations can also be observed in Spanish and French speaking children. Further, Wijnen, Krikhaar, and den Os (1994) also found this tendency for Dutch speaking toddlers.

Evidence for (pre)schoolchildren is rather limited. As mentioned above, Domahs et al. (2016) showed for written language production of German schoolchildren, that they produce writing errors in favor of rhythm and prosodic structure. For French and English preschoolers, Cavalho, Dautriche, and Christophe (2015) and Cavalho, Lidz, Tieu, Bleam, and Christophe (2016) showed that they are capable of using prosodic information to assign syntactic categories in a sentence completion task. Hearing ambiguous sentence fragments, participants were able to use prosodic information like duration from the acoustic input 'online' in order to decide whether the last word was either a verb or a noun. In a sentence repetition task, Gwinner, Gaglia, and Grijzenhout (2012) showed a tendency to repair violations of *LAPSE structures by inserting syllables for bilingual (German-Italian) preschoolers.

In sum, there is accumulating evidence for a powerful role of prosodic constraints in child language. With syntactic constraints not being fully established in child language, the influence of prosodic constraints (or a prosodic template) seems to be stronger as compared to adult language. As a result, their fulfillment may even lead to ungrammatical structures.

1.3. The interplay of ANIM and *LAPSE

McDonald et al. (1993) examined the impact of both animacy and rhythm on word order in English speech production. Adult participants were asked to memorize phrases like *doll and attic* using mental images as memory aids in order to combine the two conjuncts of a phrase. Subsequent to a variable number of mental images, participants wrote down the phrases they could remember without any order to be maintained between or within the phrases. One of the conjuncts was monosyllabic whereas the other one was either a trochee as in (5a) and (5b) or an iamb as in (5c) and (5d) (McDonald et al., 1993 [Exp. 6]). Participants were more likely to sequence the conjuncts as in (5a) and (5d) thus showing a preference for alternating rhythm and lapse-avoidance.

(5a) *dóll and áttic*

(5b) *áttic and dóll* (*LAPSE)

(5c) *dóll and antíque* (*LAPSE)

(5d) *antíque and dóll*

Interestingly, lapse avoidance was only an effective determinant for word order as long as the conjuncts didn't vary in terms of animacy, whereas animate items were produced first,

irrespective of rhythmic aspects. This preference is exemplified in (6) (McDonald et al., 1993 [Exp. 5]). When there was a combination of animate and inanimate items within a phrase, participants seemed to ignore prosodic violations (see [6]). Accordingly, the animacy constraint ANIM seems to be stronger in adults' English speech production than the rhythmic constraint *LAPSE.

(6) *children and room* (*LAPSE)

Whether the same holds for child language and German is still an open question. On the one hand, findings from speech perception attest the special importance of prosodic constraints in child language (Schröder & Höhle, 2011, Gutman, Dautriche, Crabbé, & Christophe, 2015; Morgan & Demuth, 1996) and results from speech production support the *prosodic licensing hypothesis* (Demuth, 2007). At the same time, animacy clearly affects word order in child language, too (Demuth, 2005; Prat-Sala et al., 2000; Drenhaus & Féry, 2008; Byrne & Davidson, 1985). To determine the degree to which ANIM and *LAPSE affect the choice of a particular word order, we set out to directly compare the interplay of these constraints in child versus adult speech production.

1.4. The present study

In the present study, we tested speech production by German preschoolers (Experiment 1) and adults (Experiment 2), eliciting coordinated noun phrases (*Klavier und Ratte*, 'piano and rat') from picture prompts. We systematically varied a) the animacy and b) the stress pattern of the nouns involved. The order of the conjuncts that the participants chose reveals the extent to which participants comply with the critical non-syntactic constraints on word order, namely ANIM (animate > inanimate) and *LAPSE (avoid two unstressed syllables). Experiment 3 was conducted as a control Experiment.

2. Experiment 1

We designed a picture naming task for preschool children. Participants were asked to name pairs of pictures as in **Figure 3** using conjunctive phrases without determiners (*Pilot und Klavier*, 'pilot and piano') and with no prespecified order. Pictures were coupled

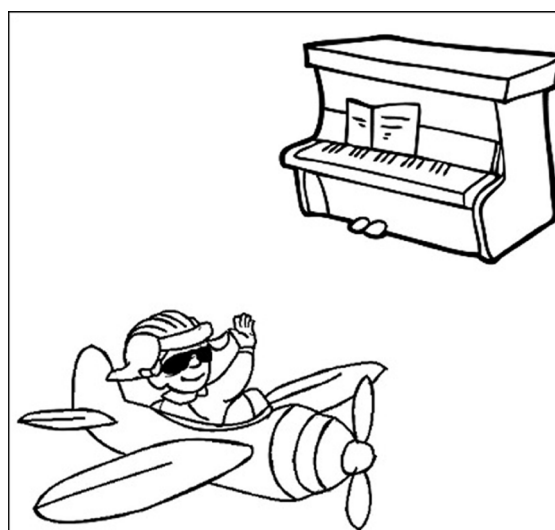


Figure 3: Stimulus with the target structure *Pilót und Klavíer* ('pilot and piano') or *Klavíer und Pilót* ('piano and pilot'). <https://www.coloringpagesfree.net/coloring-sheets-free/work-coloring-pages/pilot-invitation-colouring-sheet-free-3012.html>, <https://www.coloringpagesfree.net/coloring-sheets-free/school-colouring-pages/piano-image-to-color-462.html>.

so that response noun conjunct sequences led to violations of either ANIM (*Klavier und Rátte*, ‘piano and rat’), or *LAPSE (*Rátte und Klavier*, ‘rat and piano’), both constraints at the same time (*Hóse und Pilót*, ‘trousers and pilot’), or none (*Pilót und Hóse*, ‘pilot and trousers’). We examined the constraints’ influence on sequencing the nouns within a phrase.

Based on the hypothesis that both constraints influence word order, our prediction for the example in **Figure 3** was that the phrase *Pilót und Klavier* (‘pilot and piano’) with the animate item being sequenced first should be preferred. The reversed sequence *Klavier und Pilót* (‘piano and pilot’) would lead to a dispreferred violation of ANIM. Further, we predicted that iambs should be sequenced before trochees (*Klavier und Hóse*, ‘piano and trousers’) as a trochee in initial position followed by the unstressed conjunction necessarily leads to a violation of *LAPSE (*Hóse und Klavier*, ‘trousers and piano’). Consequently, participants had to start with an unstressed syllable to produce a rhythmic sequence, and therefore with a syllable that can’t be parsed into a trochaic foot. For English toddlers, Gerken (1994a) showed that they tend to omit (initial) unparsed syllables. In Experiment 1 we set out to test whether rhythmic alternation and animacy influence word order in German preschoolers despite such an unparsed syllable. Note that we chose disyllabic (trochaic or iambic) nouns throughout, rather than a combination of monosyllabic and disyllabic ones like in McDonald et al. (1993), to exclude an effect of word length to be responsible for a preference of rhythmic sequences.

2.1. Methods

2.1.1. Participants

Eighteen preschool children (ten female, eight male, age range 3;4 to 6;10, mean = 5;3, *SD* = 1) with typical language abilities participated in the study. The participants were recruited from regular German kindergartens in Hesse and Berlin. Children’s language abilities were normal, according to TROG-D norms (Fox, 2006), a standardized German language test for the detection of SLI (specific language impairment). All children had normal hearing ability and normal or corrected to normal vision. A parental questionnaire was used to exclude any neurological diseases, mental or intellectual disabilities, stuttering, or severe articulatory problems. The investigation was approved by the Ethics Committee of the *Deutsche Gesellschaft für Sprachwissenschaft* (DGfS, vote 15/08). For each participant, a legal guardian had signed the declaration of consent to the procedure and audio recordings.

2.1.2. Materials

Stimuli were 20 child-oriented black and white drawings, corresponding to disyllabic nouns (20 target items).⁶ We systematically crossed word stress patterns (iamb/trochee) with animacy (animate/inanimate) to yield four word-types (see **Table 1**). We used five different target items per item type. (The four word-types were then combined to form six different word/item-pairs (e.g., *pilot and feather*, see **Table 2**).

Stimuli were colouring pictures for children taken from free Internet portals (listed below). Their recognizability and their naming agreement was assessed by three independent raters. We used black and white drawings with approximately identical visual complexity and line width to balance picture pairs for visual prominence (Meyer, Roelofs, & Levelt, 2003). To minimize any systematic differences in visual prominence (Clarke, Elsner, & Rohde, 2015), three independent raters evaluated the stimuli according

⁶ Please note that due to copyright issues, most of the pictures used for the stimuli could not be included in the paper. We’ve replaced the pictures for pilot (**Figures 3 and 6**), chair, and dog (**Figure 4**) with similar ones. The original set of stimuli used in this study can be requested from the first author.

Table 1: Item types and examples.

item type	animate trochaic (a-troch)	animate iambic (a-iamb)	inanimate trochaic (i-troch)	inanimate iambic (i-iamb)
example	<i>Käfer</i> 'beetle'	<i>Pilot</i> 'pilot'	<i>Feder</i> 'feather'	<i>Klavier</i> 'piano'

Table 2: Pair types and examples (in italics) with their English translation.

pair type	a-iamb	a-troch	a-troch	a-iamb	a-iamb	i-troch
	a-troch	i-troch	i-iamb	i-troch	i-iamb	i-iamb
Examples	<i>Delfin</i>	<i>Käfer</i>	<i>Käfer</i>	<i>Delfin</i>	<i>Pilot</i>	<i>Feder</i>
	<i>Käfer</i>	<i>Feder</i>	<i>Klavier</i>	<i>Koffer</i>	<i>Klavier</i>	<i>Ballon</i>
	'dolphin'	'beetle'	'beetle'	'dolphin'	'pilot'	'feather'
	'beetle'	'feather'	'piano'	'suitcase'	'piano'	'balloon'

to their visual prominence within the corresponding pair. The three raters were instructed to allocate ten points between the two pictures of every pair so that the more prominent a picture appeared in relation to its 'partner' the more points had to be given. For example, a ratio of 4:6 meant that the picture on the left was perceived slightly less dominant. If at least two of the three raters detected such an imbalance, the more dominant picture was slightly downsized. Some of the drawings were also reworked erasing or adding single lines to make them more comparable in visual prominence and complexity.

A crucial criterion for the selection of target items was the possibility of unique graphical mapping leading to a high naming agreement. We therefore avoided target items with common synonyms. Further, all targets were on basic semantic level and morphologically simple (the only exception being *Fahrrad*, 'bike'). Items with initial consonant clusters that are typically learned after the age of 3;11 years (Fox-Boyer, 2015) were excluded. To minimize other factors than animacy and stress influencing naming sequences, item pairs were matched for frequency according to *chiLdlex* (Schroeder, Würzner, Heister, Geyken, & Kliegel, 2014), a corpus based on children's books for three different age-groups. With age group six to eight years being closest to our preschool participants, frequency data were taken from this sub-database. In cases of a (small) frequency difference, item pairs were combined in a way that frequency would work against our hypothesis, e.g., if a pair of an animate and an inanimate item differed in frequency, the latter was chosen to be more frequent. Furthermore, vowel height was considered as a factor since high front vowels are known to be preferably produced before low vowels (e.g., [i] > [a]) (Lohmann, 2013). With [i] as the highest and [a] as the lowest German vowel, combinations of these two as nuclei of stressed syllables were excluded (as it would be the case in a combination of *Pirát*, 'pirat' and *Delfín*, 'dolphin'). Moreover, we made sure that there was no direct semantic relation between the items or within a pair. Finally, pairs appearing in idioms were excluded.

2.1.3. Stimulus presentation

Stimulus picture pairs consisted of drawings of two different item types each, yielding six different pair types as shown in **Table 2**. Every single item was presented within three different pairs (the complete set of stimuli is listed in the Appendix). We used five different item pairs per type yielding 30 different combinations.

Picture pairs were presented in a squared frame (see **Figures 3** and **4**), separated by an invisible diagonal line running from top left to bottom right. Pictures did not surpass the diagonal so that they were clearly located bottom left and top right within the square.

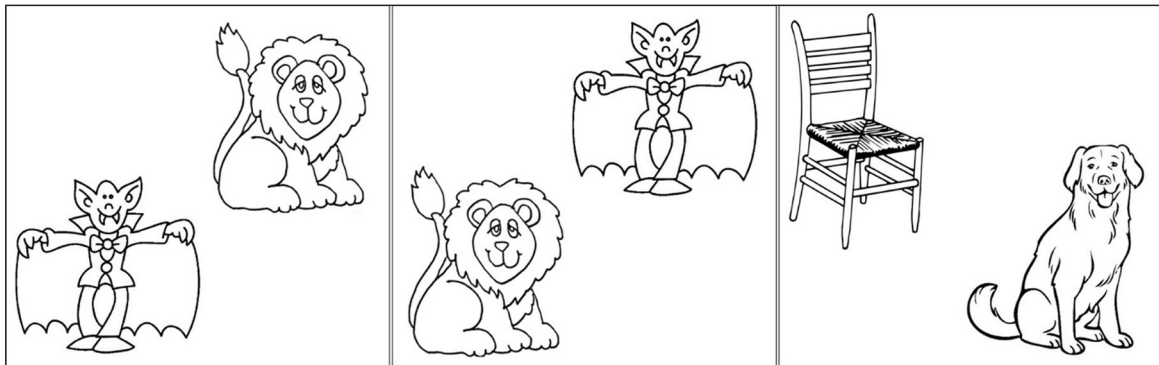


Figure 4: Stimulus arrangements in item pair pictures. The left and middle panel show examples of an experimental item pair in the two spatial orders and the right panel shows a filler pair. https://www.ausmalen2000.com/img.php?id_img=6341, https://www.ausmalen2000.com/img.php?id_img=7384, <https://openclipart.org/detail/276556/chair-3>, <https://www.bestausmalbilder.de/ausmalbild/labrador-hund/>.

We chose this diagonal arrangement to counteract the tendency of left to right or top to bottom naming (Knudson, Fischer, & Aschersleben, 2014). Further, picture pairs were also presented in reversed spatial order (**Figure 4**, left and middle panel) so that every pair was used twice yielding 60 item pairs altogether. They were divided into two pseudo randomized lists so that every combination (either the original item pair or its reversed version) appeared exactly once in a list.

We added 16 filler pairs arranged in the opposite direction to the item pairs. Thus, filler item pictures were placed top left or bottom right inside the square with the invisible diagonal line running from bottom left to top right (**Figure 4**, right panel). These fillers were used to minimize automatized naming sequences (e.g., naming always the top right picture first) that could result from an invariant visual structure. All filler pairs were added to both lists yielding a total of 46 picture pairs (30 experimental and 16 filler pairs) per list. Additionally, filler pairs were combined in a list of familiarization items, where they were arranged in all possible spatial orders (bottom left and top right as well as top left and bottom right).

2.1.4. Procedure

For every child there was one experimental session preceded by some familiarization with the material. The familiarization was conducted to ensure that picture names were as consistent as possible across participants. It also served to activate the participants' representation of the names to a high degree, with the expectation that, in this way, potential effects of frequency or recency⁷ (Narasimhan & Dimroth, 2008, 2017) would be weakened or leveled out. The first familiarization phase was conducted by the parents at home using a colouring book including the target pictures and corresponding names. The book was sent to the family at least three days prior to the actual investigation. The experimental session was conducted in a quiet room in the kindergarten or at the family's home. Participants got a short and child-oriented explanation concerning the general purpose of the investigation ('exploration of children's speech') without providing details about the specific aim of the experiment. One session took 30–45 minutes. Pauses and games were inserted individually depending on the children's needs. **Table 3** illustrates the whole procedure.

⁷ Studies by Narasimhan and Dimroth (2008, 2017) reported effects of recency or newness on the order of conjuncts. Specifically, they found that children tend to name newly introduced referents before old and familiar referents whereas adults tend to name familiar referents before new ones. However, at least in children, the effect appears to be unstable (DeRuijter, Narasimhan, Chen, & Lack, 2018).

Table 3: Overview of the whole procedure.

	preparation	experimental session			
Phase	familiarization items	familiarization items	familiarization target structure	experiment	language assessment
Materials	colouring book	memory game	filler list	list 1 or 2	Trog-D
Time	3 days prior to investigation	10 minutes	5 minutes	10 minutes	10–20 minutes

As can be seen in **Table 3**, a first familiarization took place in advance and a second and third during the actual experimental session. For the first familiarization children got the colouring book *Sprachforscherbuch* ('book for speech/language scientists') which was purpose-made for the study. It contained instructions for the parents as well as all 20 stimulus pictures with their corresponding target-words underneath. Parents were instructed to look at the book together with their child and see if the child named the pictures correctly by herself. If not so, parents were asked to make sure that their child heard every correct target-word three times and that he or she named every picture at least once by herself using the correct target-word to further increase naming agreement for the target pictures. Children could colour the pictures as they liked.

The second familiarization with the materials was done with the investigator during the experimental session using a memory game, which contained cards with all stimuli pictures. Rules were similar as in an ordinary memory game with every person revealing two cards in one turn, whereas one of those cards could be left open. Once a player discovered two identical pictures, he or she had to clap one of the corresponding cards for keeping the pair. During the game, all 20 pictures were named by the child at least once and the experimenter already modeled the target structure (i.e., a conjoined noun phrase) when opening two cards.

The third familiarization was done to make sure that participants had understood the experimental task. The children sat in front of a screen (Samsung Notebook 900X, model: NP900X3c) on which the pairs of filler items were presented. Every participant got the same list of familiarization items. Children were instructed to name the pictures within a conjoined noun phrase without determiners. If children asked which sequence they should use, they were told that it didn't matter. For the younger children, the task was demonstrated by the experimenter using two hand-dolls, a wizard, and a pirate. The wizard explained the task and modeled the target structure (e.g., *Klavier und Planet*, 'piano and planet'). After some examples, the pirate named the picture pairs by himself, but he did not understand the task and made mistakes. Participants were encouraged to help the pirate and to name the item pairs for him. During this phase, the wizard corrected the children and explained the task again, if necessary. The older children (usually the six-year-old) got the instruction without hand puppets. In both cases, the experimenter named the picture pairs in a way so that the first named picture varied in its spatial location. Every target was named in individual speed so that the picture pairs on the screen changed as soon as the participant or the experimenter had named them. The third familiarization phase ended when the participant had named at least three subsequent picture pairs correctly in a conjoined noun phrase. Before the actual experiment started, children were familiarized with the digital recorder (Olympus VN-8700OC). Their responses were recorded throughout the experiment.

The two experimental lists (Appendix) were randomly assigned to participants so that nine children got list one and the other nine list two. The procedure for the experimental

lists was the same as in the third familiarization, whereas the instructions were usually shorter. During the experimental session, the experimenter did not interrupt the participant but gave corrective feedback when the child didn't produce a correct target structure (irrespective of its order).

2.1.5. Analyses

Prior to analysis, all recordings were transcribed by an undergraduate student who was naïve to the purpose of the study. Utterances were classified for their validity and noun sequence. Only valid responses were entered into the analysis. An utterance was considered valid if it contained the correct target words (i.e., bare nouns) as well as the conjunction *und* ('and') in between. Determiners or any other linguistic material invalidated the response. Filled pauses (e.g., *ehm*) also lead to an exclusion of the whole phrase from analyses. Noun sequences were coded as a change or maintenance of an arbitrarily predefined standard word order. This standard word order was defined as the bottom left picture preceding the top right picture. The alphabetical order in **Figure 5** visualizes the predefined standard sequence.

Note that this 'standard sequence' was arbitrarily set, i.e., we did not predict participants to name A preferably before B or vice versa. Rather, we predicted that, due to the reported top > bottom and left > right bias (Knudson et al., 2014), there should be some variation in fronting the bottom left A or top right B. Hence, the maintenance or change of the arbitrarily defined standard sequence $A > B$ was our dependent variable. This decision was also based on the fact that invalid responses might lead to an unbalanced number of data points per condition: Since the spatial arrangement of pictures was balanced throughout the conditions, a loss of data points could cause an apparent effect of position to be mistaken for an effect of our independent variables—at least in a coding that is unrelated to position.

The actually produced sequence that either adhered to, or deviated from, the standard sequence was taken as the dependent variable for the evaluation of the effects of ANIM and *LAPSE (see **Figure 6**). This binary response variable was coded as zero (0) if picture A (bottom left) was named before B (top right)—that is in the case of adherence to the standard sequence—and as one (1) when the production deviated from this standard sequence (i.e., when the noun in position B was the first conjunct named in the phrase).

Animacy, and prosody (i.e., word stress) were taken as predictors for maintenance of, or deviation from, the standard sequence. Depending on the item combination in a pair, we coded each predictor as favoring (.5) or disfavoring (-.5) a deviation from the predefined standard sequence. Zero (0) was coded for cases in which the predictor in question did not make a difference (that is, when the two target nouns had the same value for the respective predictor). The participants' age and the difference of the logarithmized word frequencies

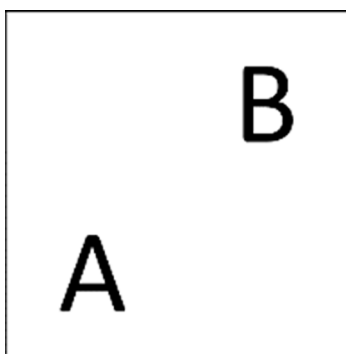


Figure 5: Schematic depiction of a stimulus pair with $A > B$ as predefined standard sequence.



Hypothesis and prediction	Example stimulus and standard sequence	Example coding and prediction
<p>ANIM is effective.</p> <p><i>Animates are produced before inanimates.</i></p>	 <p>Picture pair for standard sequence: <i>Pilót und Klavier</i>, 'pilot and piano' (animate/iambic & inanimate/iambic)</p>	<p>animacy: -.5 rhythm: 0</p> <p>Prediction: adherence to standard sequence <i>Pilót und Klavier</i>, 'pilot and piano' (animate/iambic & inanimate/iambic)</p>
<p>*LAPSE is effective.</p> <p><i>Iambs are produced before trochees.</i></p>	 <p>Picture pair for standard sequence: <i>Lówe und Vampír</i>, 'lion and vampire' (animate/trochaic & animate/iambic)</p>	<p>animacy: 0 rhythm: .5</p> <p>Prediction: deviation from standard sequence <i>Vampír und Lówe</i>, 'vampire and lion' (animate/iambic & animate/trochaic)</p>

Figure 6: Illustration of the hypotheses, their predictions (in italics), and associated codings. The upper row illustrates our predictions for ANIM working as a constraint with animacy as the predicting factor: animate items (e.g., *Pilót*, 'pilot') are expected to be sequenced before inanimate ones (e.g., *Klavier*, 'piano'). The bottom row shows our prediction for *LAPSE working as a constraint with stress (or rhythm) being the predicting factor: To avoid lapses (*Lówe und Vampír*, 'lion and vampire') iambs should be sequenced before trochees (*Vampír und Lówe*, 'vampire and lion'). Both factors are coded with respect to the arbitrarily predefined standard sequence (middle column) where the picture on the bottom left is named before the one on top right. In the example for ANIM (upper row), animacy is reducing the probability of changing the standard sequence (coded as -.5) as the bottom-left picture is animate (*Pilót*, 'pilot') and the top-right picture is inanimate (*Klavier*, 'piano'). Stress does not influence the sequence (coded as 0) as both pictures have referents with the same stress pattern. In the example for *LAPSE (lower row) animacy does not influence word order (coded as 0) as both pictures are animate. The factor stress enhances the probability of changing the standard sequence (coded as .5) as it yields a structure inducing a stress lapse (*Lówe und Vampír*, 'lion and vampire') whereas reversing the standard sequence yields a rhythmic sequence (*Vampír und Lówe*, 'vampire and lion'). <https://www.coloringpagesfree.net/coloring-sheets-free/work-coloring-pages/pilot-invitation-colouring-sheet-free-3012.html>, <https://www.coloringpagesfree.net/coloring-sheets-free/school-colouring-pages/piano-image-to-color-462.html>, https://www.ausmalen2000.com/img.php?id_img=7384, https://www.ausmalen2000.com/img.php?id_img=6341.

between noun A and noun B of each pair were entered as covariates. In order to observe the animacy constraint (ANIM), animate nouns should be produced before inanimate ones in more than 50% of the pairs. The rhythmic constraint favors the sequence *iamb “und” trochee*.

2.2. Results

On average, each of the 18 participants produced ~21 valid conjoined noun phrases (range: 11 to 28) on the basis of 30 stimuli pictures. This corresponds to 385 valid phrases obtained in the experiment. Interestingly, in the subset with both nouns being iambs, out of the 90 produced noun phrases 48 were valid (53.3%), while in the subset with both nouns being trochees, 75 of the 90 recorded phrases were valid (83.3%). In the remaining

mixed subset (containing iambs and trochees in each pair) out of 360 noun phrases 262 were valid (72.8%). In the following description the ratio between two manifestations of each factor is used to provide a summary of the descriptive statistics, followed by two mixed effect models with the full data set and a subset.

Figure 7 summarizes the results with respect to the following factors: Position of the items (that is the visual stimulus position on the screen), word frequency, animacy, and prosody and their influence on conjoined phrase ordering in Experiment 1. In 61% of cases, children named the top right conjunct before the bottom left one, indicating a general preference for this order. However, no child used one of the orders exclusively. Thirteen of the 18 children started the conjoined noun phrase with the top right item in more than 50% of their utterances; the remaining five children generally preferred the reverse order. Word frequency of the individual nouns did not appear to systematically affect conjunct ordering, as could be expected given that item pairs were approximately matched for word frequency. Overall, the more frequent noun was named first in 48% of cases, as can be seen in **Figure 7**. The factor animacy clearly affected word order. For the subset of stimuli in which the nouns differed in terms of animacy, the animate noun was named first in 65% of cases (see **Figure 7**).

Examining all pairs with rhythmically varying nouns ($n = 262$), both orders are equally present: 52% of all item-pairs with mixed word rhythm led to rhythmically alternating realizations (iamb > trochee; see **Figure 7**). However, when studying the subset of data in which there is no animacy distinction between the two depicted objects (either both pictures animate or both inanimate, $n = 128$ coordinate phrases), a preference for rhythmically alternating realizations is revealed (59%).

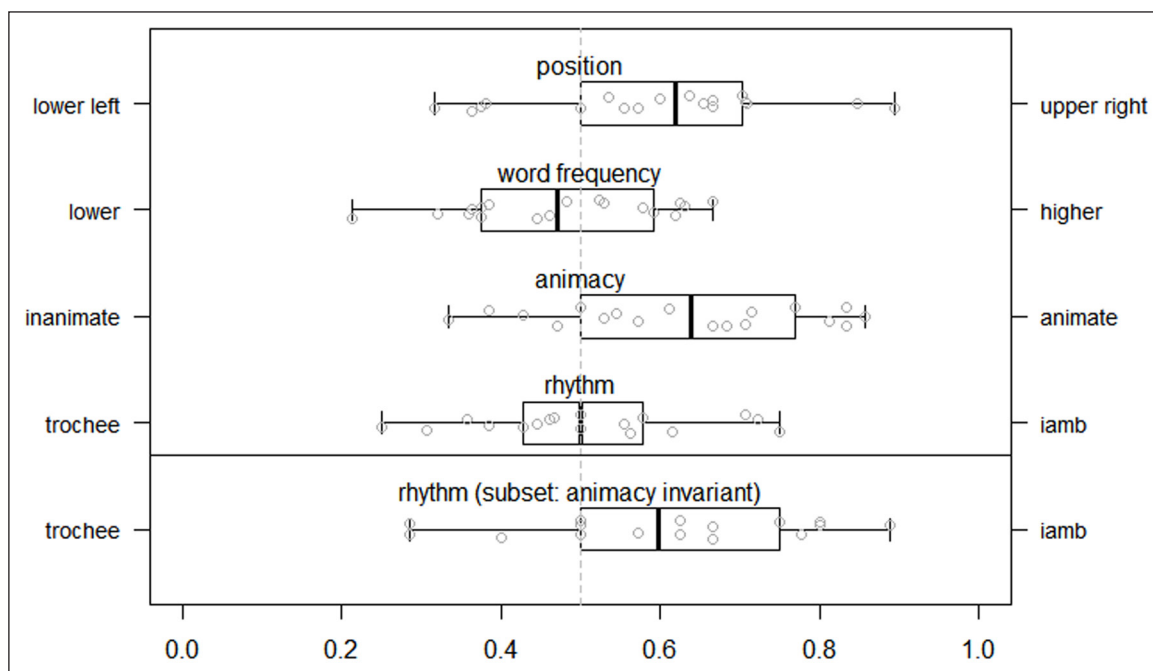


Figure 7: Mean ratio of conjunct order in Experiment 1 (children, conjunction with *und*, ‘and’) with respect to the following factors: Position (proportion upper right first), word frequency (proportion higher frequent first), animacy (proportion animate first), and prosody (proportion iamb first). Each plot represents the influence of one factor on word order. The bold line within each box indicates the median across participants. Grey circles represent means of individual participants. Chance distribution (.5 ratio) is indicated by the vertical dotted line. The four upper plots refer to the dataset where the relevant factors vary whereas the lowest plot refers to the subset where animacy doesn’t vary in the item pairs.

We also assessed whether the age of the children matters for the ordering of conjuncts. To this end, we studied the correlations of age and the other variables of interest. For each participant, we calculated the proportion of responses that a) adhere to the predetermined standard order, b) observe the animacy constraint (in item pairs that vary with respect to animacy), and c) abide by the principle of rhythmic alternation. We then correlated these proportions with the participant's age (see **Figure 8**).

The proportion of A > B orders does not appear to vary systematically with age ($r = -0.088$, $p = .73$, see the scatterplot in the Appendix). However, the strength of the observed animacy constraint (animate > inanimate) clearly increases with age. This is shown in **Figure 8** depicting a moderate positive correlation between age and preference for the animate > inanimate order ($r = .55$, $p = .019$). No significant developmental pattern is observable for the rhythmic constraint ($r = .21$, $p = .41$, see the scatterplot in the Appendix).

We employed a generalized linear mixed-effects regression model (glmer, Bates, Mächler, Bolker, & Walker, 2015) in R statistical software (version 4.0.2, R Core Team, 2020) to evaluate the main effects of ANIM and *LAPSE and their interaction on the ordering of the conjuncts. Apart from ANIM and *LAPSE, we also included age and the interactions between a) age and ANIM, and, for completeness, b) age and *LAPSE. Age was counted in months, centered to the mean age of all participants (64.4 months) and rescaled to year (division by 12). Given the findings by McDonald et al. (1993), we especially expected *LAPSE to affect ordering if ANIM is mute, i.e., when both pictures have the

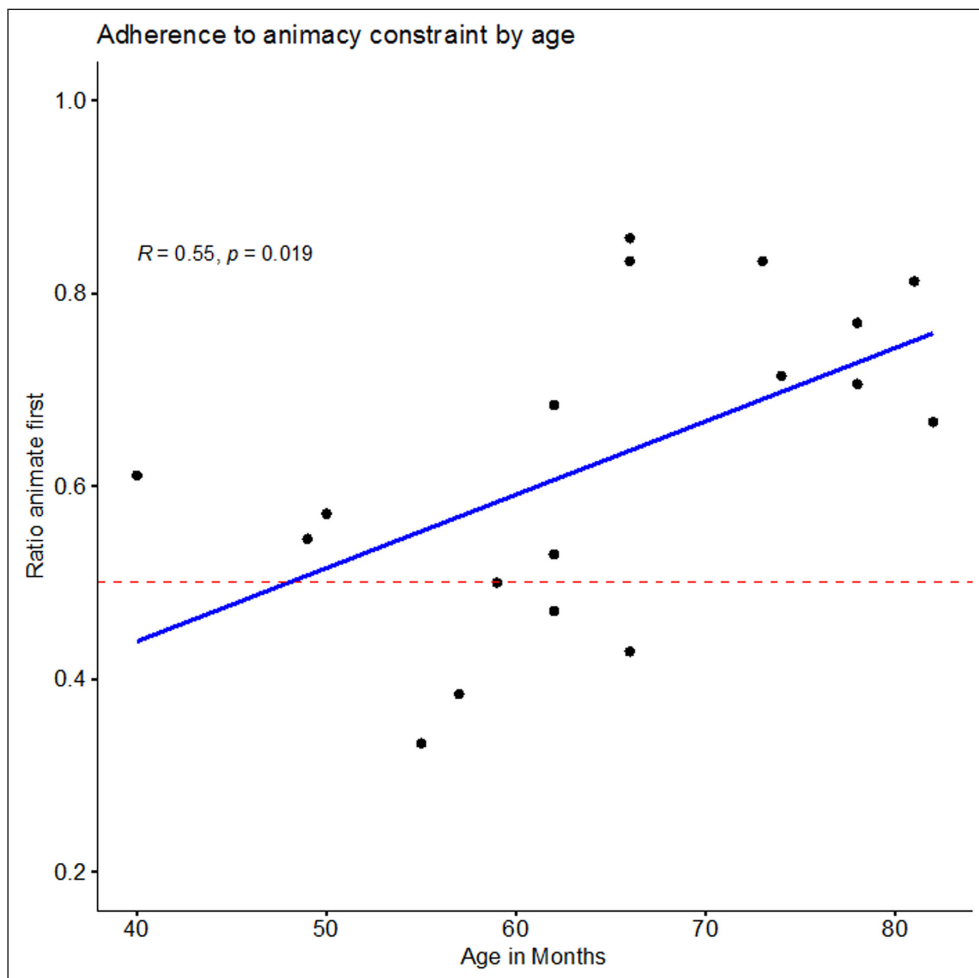


Figure 8: Scatterplot showing the proportion of animate-first responses as a function of age. The slope of the blue regression line corresponds to the correlation coefficient R.

same animacy value. We therefore introduce ANIM_Diff (coded as +.5 when the pictures differ in terms of animacy, otherwise -.5), and included its interaction with *LAPSE in the model. Finally, the effect of word frequency was included, coded as the difference in frequency (log) between the two depicted conjuncts. To account for the variability due to individual participants and individual item pairs, participant and item were included as random intercepts. In the models we report below, we did not include random slopes for the fixed effects, because the model failed to converge.

The model confirms a significant effect of ANIM. Transforming the corresponding coefficient of 1.53 (logit) to odds ratio (4.6) by exponentiation, we can interpret this coefficient in the following way: Deviating from the standard order (A > B in **Figure 5**) in favor of the animacy constraint is nearly 5 times as likely as deviating from it to the detriment of the animacy constraint. The significant interaction of ANIM and age (logit: .861; odds ratio: 2.37) reflects the above-mentioned correlation (see **Figure 8**) and suggests that (for the age range studied here) aging by 12 months makes it roughly twice as likely to observe the animacy constraint. The significant *LAPSE:ANIM_Diff interaction confirms that adhering to the rhythmic constraint (logit -1.876) is roughly 6 times as likely when the pictures do not vary in terms of animacy compared to when there is an animacy difference between them. All other main effects and interactions do not reach conventional levels of significance (see **Table 4**).

Given the *LAPSE:ANIM_Diff interaction, we tested the effect of *LAPSE specifically in the subset in which item pairs did not vary with respect to animacy. In this subset, as shown in **Figure 7**, participants were more likely to follow the rhythmically alternating order *iamb* “und” *trochee*. The corresponding generalized linear model for this subset included the same factors as the full model, except for the predictors related to animacy and the corresponding interactions. The model statistics are summarized in **Table 5**.

The significant main effect of *LAPSE (logit .83 – odds ratio 2.29) confirms that, when faced with item pairs that do not vary in terms of animacy, participants were roughly twice as likely to deviate from the predetermined standard order in favor of a rhythmically alternating rendition than to deviate from it to the detriment of rhythmic alternation.

2.3. Discussion of Experiment 1

In Experiment 1 we set out to test whether animacy (ANIM) and rhythm (*LAPSE) have an impact on the ordering of nouns in conjoined phrases in the speech production of German preschoolers. For this purpose, we designed a picture naming study: Participants

Table 4: Results of the generalized linear mixed-effects regression model (glmer) for the whole data set with rhythm, animacy, age, their interactions, and word frequency as fixed effects.

	Estimate	SE	z value	p value
(Intercept)	0.53636	0.22043	2.433	0.01496*
ANIM	1.53279	0.31515	4.864	<0.001***
*LAPSE	-0.02010	0.30316	-0.066	0.94713
Age	0.02522	0.19392	0.130	0.89651
ANIM_Diff	-0.16994	0.33877	-0.502	0.61593
Freq_Diff_log	-0.25905	0.27327	-0.948	0.34316
ANIM:*LAPSE	1.21218	1.11107	1.091	0.27527
ANIM:Age	0.86098	0.31211	2.759	0.00580**
*LAPSE:Age	0.21557	0.30117	0.716	0.47414
*LAPSE:ANIM_Diff	-1.87605	0.59260	-3.166	0.00155**

Table 5: Results of the generalized linear mixed-effects regression model (glmer) for the subset in which the two nouns are either both animate or both inanimate. Fixed effects are rhythm, age, their interaction, and word frequency.

	Estimate	SE	z value	p value
LAPSE	0.829543	0.403483	2.056	0.03979
Age	0.005334	0.016895	0.316	0.75223
Freq_Diff_log	-0.362037	0.452736	-0.800	0.42390
*LAPSE:Age	0.016924	0.033854	0.500	0.61714

were asked to name pairs of pictures using conjunctive phrases without determiners (e.g., *Pilot und Klavier*, ‘pilot and piano’) and with no prespecified order. Pictures were coupled and presented diagonally in a square with one picture being placed lower left and the other upper right. Due to the design (we used animate and inanimate items that were either trochaic or iambic) the chosen sequences yielded rhythmically alternating or disrhythmic phrases which either complied with the animacy constraint or violated it.

The results show that the preschoolers as a group prefer to start with the upper right picture (i.e., deviate from the arbitrarily set standard sequence). This is consistent with a study by Knudson et al. (2014) who found that the left-right bias (that is commensurate with the reading direction) only comes into effect after school enrollment. Next to this general effect of presentation layout, we found a strong effect of animacy favoring animate referents being named first in line with findings reported by Prat-Sala et al. (2000). Additionally, the effect of animacy correlated positively with the age of the children. Since the age range of our participants is rather large, especially in relation to the relatively small sample size, the results need to be interpreted with caution. However, the clear age effect of ANIM is noteworthy. Interestingly, this result is consistent with the possible reasons for animacy effects mentioned in the introduction: For example, the account of ‘concept accessibility’ which was proposed by Bock and Warren (1985) assumes the higher number of semantic pathways leading to an ‘animate’ lexical entry to accelerate word access. Since the lexical entries are getting richer and more diverse during language development, it seems plausible that an effect of animacy due to concept accessibility increases with age.

The effect of rhythm was weaker and only significant for the subset of item pairs in which animacy was held constant. The rather weak effect of *LAPSE, our rhythmic constraint, might be due to the fact that in order to produce a rhythmic sequence (*Pilót und Rátte*, ‘pilot and rat’), the children had to start with an unstressed syllable (*Pi-* in this example). According to Gerken (1994a), toddlers avoid initial unstressed syllables since they cannot be parsed into a foot. It is possible that this avoidance of unparsed initial syllables (or the higher accessibility of trochees as opposed to iambs, see Schiller, Fikkert, & Levelt, 2004) affected our results and thus weakened the effect of *LAPSE. The higher number of invalid data points with iambs as opposed to trochees supports this explanation. Nevertheless, an effect of rhythm on conjunct order is effective in spite of these potential counteracting forces. In the next experiment we set out to test how German adults sequence bare nouns in conjoined phrases and how the discussed constraints affect their choices.

3. Experiment 2

3.1. Methods

3.1.1. Participants

Participants were 34 young adults (age range 20 to 37, mean = 23;7, *SD* = 3;9) with German as (one of) their native language(s). Gender among participants was roughly

balanced (20 females, 14 males). All participants were students at the Goethe University of Frankfurt, Germany.

3.1.2. Materials, procedure, and analysis

The materials and procedure of Experiment 2 were almost identical to Experiment 1 with the main difference being a shorter familiarization and instruction, and additional filler tasks during the experiment (after every third picture pair, the participant had to give an answer to a simple arithmetic task). In order to familiarize participants with the lexical material, they were asked to look at a piece of paper showing the 20 pictures and their corresponding names in order to memorize the exact word for every picture. When they indicated that they managed to memorize the corresponding names, the piece of paper was removed and replaced by a second one that missed the corresponding names. Then, the participants had to name each picture using the words they memorized from before. In cases where they produced at least one of the corresponding names incorrectly, participants were asked to memorize them again using the first piece of paper and naming them again using the second. Once all 20 pictures were named correctly in a row, familiarization was finished, and the actual experiment started.

As in Experiment 1, the participants sat in front of a screen and were instructed to name the picture pairs in a conjoined phrase. Both lists were preceded by the same three filler pairs. These were used to illustrate the task. The presentation of one of the two lists started directly, so that while showing the first filler pair, the experimenter gave the instruction to name the picture pairs in a swift way without using determiners.⁸

In cases where a participant did not fully understand the instruction, it was repeated with reference to the next pair of fillers. Participants' responses were recorded. The two experimental lists were randomly assigned to participants so that half of the participants received list one and the other half list two. During the experimental session, the experimenter did not interrupt the participant but gave corrective feedback when the participant didn't produce a correct target structure. The analysis of Experiment 2 was identical to that of Experiment 1, apart from the fact that we did not include age as a predictor in the model.

3.2. Results

Two participants used one of the orders exclusively (either always $A > B$ or always $B > A$, cf. **Figure 5** above). Due to their lack of variation, these responses are uninformative for our hypothesis and were therefore excluded from further analysis. From the remaining participants, we obtained 950 valid responses. **Figure 9** summarizes the results with respect to the factors position (that is the visual stimulus position on the screen), word frequency, animacy, prosody, and their influence on conjunct ordering in Experiment 2.

The conjunct presented in the bottom left was named first in 65% of cases, indicating a general preference for this order. Word frequency of the individual nouns did not appear to systematically affect ordering of the conjuncts. The factor animacy (ANIM) does appear to bias word order: 58% of the animate items served were named first in the phrase (see **Figure 9**). However, the data don't suggest a clear effect of rhythm (*LAPSE): 51% of the iambic items were mentioned before the trochaic ones. When studying the subset of data in which there is no animacy distinction between the two presented objects (both animate

⁸ The experimenter gave the following instruction: "On this screen you will see picture pairs which you should name and simple arithmetic tasks for which we ask you to report the solution. You will recognize some of the pictures—in this case please use the names you memorized earlier. Please name each picture pair in a swift way with isolated nouns and an *und* ('and') in between, for example: *Papagei und Pirat* ('parrot and pirat'). Don't use any determiners like *der/die/das* ('the') or *ein/eine* ('a'). Try not to name all the picture pairs in the same spatial order but choose the sequence that sounds most natural to you."

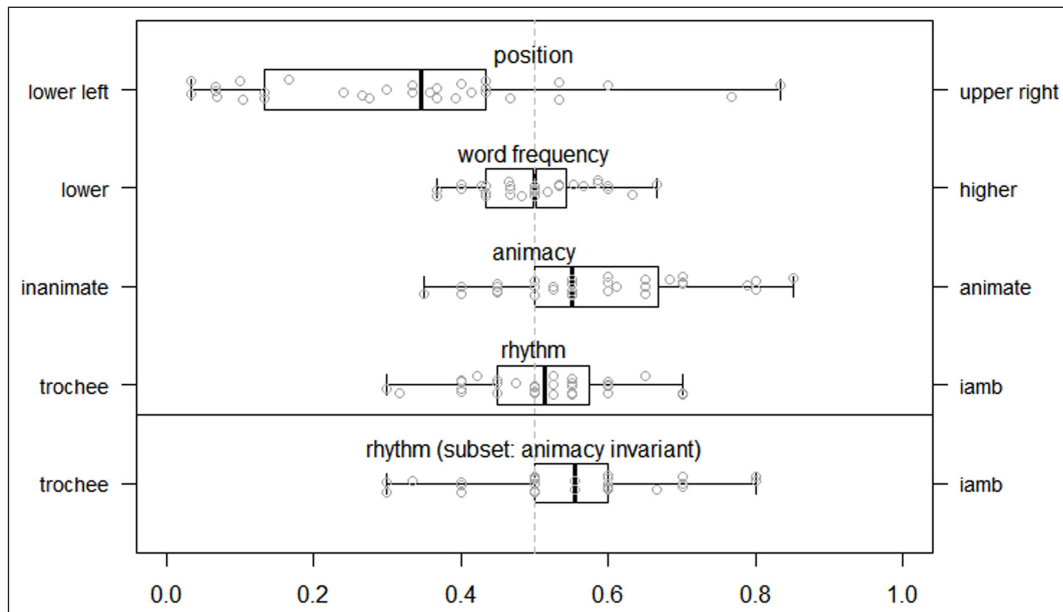


Figure 9: Mean ratio of conjunct order in Experiment 2 (adults, conjunction with *und*, ‘and’) with respect to the following factors: Position (proportion upper right first), word frequency (proportion higher frequent first), animacy (proportion animate first), and prosody (proportion iamb first). Each plot represents the influence of one factor on word order. The bold line within each box indicates the median across participants. Grey circles represent means of individual participants. Chance distribution (.5 ratio) is indicated by the vertical dotted line. The four upper plots refer to the dataset where the relevant factors vary whereas the lowest plot refers to the subset where animacy doesn’t vary in the item pairs.

or both inanimate, $n = 316$ coordinate phrases), the distribution of responses shows a preference for rhythmic sequences (in 55% of cases the iambic noun was named first).

A generalized linear mixed-effects regression model (glmer, Bates et al., 2015) was employed in R statistical software (version 4.0.2, R Core Team, 2020) to evaluate the effects of ANIM, *LAPSE, their interaction, the effect of uniformity versus difference regarding the animacy of the depicted objects (ANIM_Diff), and its interaction with *LAPSE, as well as word frequency on word order. To account for the variability due to individual participants and items, these were included as random intercepts (we did not include random slopes due to convergence errors). Apart from the significant intercept (reflecting the clear preference for the left before-right order), the model revealed a significant effect of ANIM (logit .819 – odds ratio ~ 2.27), suggesting that deviating from the standard sequence in favor of ANIM is roughly twice as likely as deviating from the standard order to the detriment of ANIM. Also, the significant interaction of *LAPSE and ANIM_Diff suggests that any effect of rhythm is contingent upon whether or not the depicted objects vary in terms of animacy. All other main effects and the interaction did not significantly affect conjunct ordering. In a second step, we isolated animacy-invariant item pairs. In this subset, there was a significant effect of *LAPSE (logit: .656 – odds ratio: 1.93), suggesting that deviation from the standard order was nearly twice as likely when this served rhythmic alternation in the response. The model statistics for both the full data set and for the subset are summarized in **Table 6**.

3.3. Discussion of Experiment 2

Experiment 2 was a replication of Experiment 1 conducted with young adult participants. In line with Experiment 1, we found a clear influence of ANIM on ordering conjuncts and a significant impact of *LAPSE on conjunct order for the subset of stimuli that do not vary in terms of animacy. In contrast to the preschoolers, the group of adult participants

strongly preferred to name the items placed in the lower left quadrant first, which is in line with Knudson et al. (2014) who showed that the left-right bias emerges with reading proficiency. This relatively strong bias regarding the dependent variable may curb the other effects and thus explain the considerably weaker effect of ANIM in the group of adult participants (odds ratio ~2) when compared to the preschoolers (odds ratio ~4). The effect of *LAPSE in the subset analysis, however, was more comparable in the two experiments (odds ratios in both groups are around ~2).

In order to produce a rhythmic sequence, participants had to name iambic nouns before trochaic ones. Since Schiller et al. (2004) found shorter naming latencies for trochees, rhythm needs to counteract this tendency when favoring a fronted iamb. This might have weakened the results, similarly to the earlier mentioned prosodic issues with the unparsed initial syllable in our rhythmic structures. Again, we would like to highlight the fact that despite this potentially counteracting effect of word form retrieval, the rhythmic constraint *LAPSE *did* influence word order.

In order to shed light on the question whether latencies of word form retrieval might have influenced our results, we conducted Experiment 3 as a control Experiment. Further, by exchanging *und*, ‘and,’ with *oder*, ‘or,’ the target structure yields another deviation from alternating rhythm, i.e., a stress clash (two or more stressed syllables in a row), as illustrated in Figure 10.

Table 6: Results of the generalized linear mixed-effects regression model (glmer) with rhythm, animacy, their interaction, and word frequency as fixed effects. The top part shows the results for the whole data set; the bottom part shows the results for the subset of item pairs in which the two nouns are either both animate or both inanimate.

Model (glmer) for complete data set				
	Estimate	SE	z value	p value
(Intercept)	-0.893038	0.203613	4.386	<0.001***
ANIM	0.830116	0.189704	4.376	<0.001***
*LAPSE	0.172197	0.192992	0.892	0.3723
ANIM_Diff	-0.003756	0.161272	-0.023	0.9814
Freq_Diff_log	0.179736	0.169181	1.062	0.2881
ANIM:*LAPSE	0.584501	0.530518	1.102	0.2706
LAPSE:ANIM_Diff	-0.791485	0.37473	2.112	-0.0347
Model (glmer) for animacy-invariant subset				
	Estimate	SE	z value	p value
(Intercept)	-0.925	0.253	-3.654	<0.001***
LAPSE	0.656	0.285	2.304	0.021
Frequency	0.382	0.327	1.171	0.242

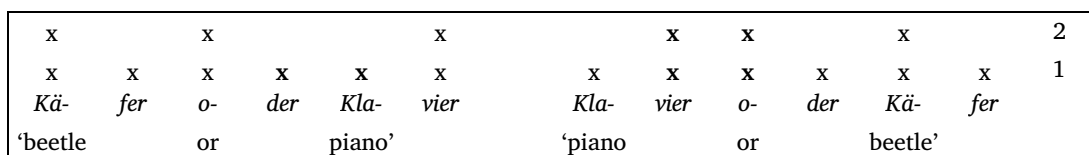


Figure 10: Simplified metrical grid illustrating a stress lapse (left panel) and a stress clash (right panel). The violations of *LAPSE and *CLASH are bold-faced. Layer 1 represents the default beat for every syllable. Layer 2 represents the beat for every syllable bearing lexical stress.

4. Experiment 3

In this final experiment, participants named the same stimuli as in Experiment 1 and 2, this time using the disjunction *oder* ('or') in between (e.g., *Klavier oder Käfer*, 'piano or bug'). This disyllabic disjunction is trochaic, hence with *oder* ('or') the formerly rhythmic sequence with a fronted iamb (*Klavier und Käfer*, 'piano and beetle'), now yields a stress clash with the two adjacent syllables *-vier* and *o-* both being stressed (*Klavier óder Käfer*, 'piano or beetle'). Conversely, the structure with the fronted trochee (*Käfer und Klavier*, 'beetle and piano') now yields a stress lapse with the two adjacent syllables *-der* and *Kla-* both being unstressed (*Käfer óder Klavier*, 'beetle or piano,' see **Figure 10**). Thus, in Experiment 3 both ways of sequencing iambs and trochees are dysrhythmic—note that if differences in lexical access would have caused the rather weak effect of *LAPSE in Experiments 1 and 2, then this should be replicated in Experiment 3 given that the same stimuli are used.

4.1. Methods

4.1.1. Participants, materials, procedure, and analysis

Twenty-eight of the participants from Experiment 2 participated in this follow-up study. Materials and procedure of Experiment 3 were identical to Experiment 2 with two exceptions: First, in Experiment 3 participants had to name the picture pairs with an *oder*, ('or') as conjunction (as opposed to the *und*, 'and' in Experiment 2). Second, since Experiment 3 directly followed Experiment 2 (with a pause of approximately 5 minutes), each participant was presented the very list that he or she had not seen in Experiment 2. The analysis of Experiment 3 was identical to Experiments 1 and 2.

4.2. Results

On average, each of the 28 participants produced 29.6 valid conjoined noun phrases (range: 27 to 30) on the basis of 30 stimulus pairs. This yielded 857 valid phrases obtained in the experiment. **Figure 11** summarizes the results with respect to the following factors: position, word frequency, animacy, and prosody and their influence on conjoined phrase ordering in Experiment 3.

In 65% of the cases, participants named the bottom left conjunct before the top right one, indicating a clear general preference for this order. One participant used one of the orders (lower left > upper right) exclusively and was therefore excluded from further analysis. Twenty-four of the 27 remaining participants started the conjoined noun phrase with the referent in the bottom left picture in more than 50% of their utterances.

The factor animacy (ANIM) affected word order in the whole dataset: 54% of the animate items were sequenced before the inanimate ones (see **Figure 11**). Again, the data don't show a clear effect of rhythm (*LAPSE, 51% of the iambic items were named before the trochaic ones). When studying the subset of data in which there is no animacy distinction between the two presented objects (both animate or both inanimate, $n = 282$ coordinate phrases), no clear preference for naming trochees or iambs first is revealed either (52% iambs first).

A generalized linear mixed-effects regression model (glmer, Bates et al., 2015) was employed in R statistical software (version 4.0.2, R Core Team, 2020) to evaluate the effects of ANIM and *LAPSE, their interaction, as well as of logarithmized word frequency on word order. To account for the variability due to individual participants and items, these were included as random intercepts. Apart from the significant intercept (reflecting the clear preference of left before right-order), the model reveals a significant effect of ANIM. Transformed to odds ratio (1.56), this effect appears to be somewhat weaker than what was observed in Experiment 2. Even though frequency does not seem to affect word

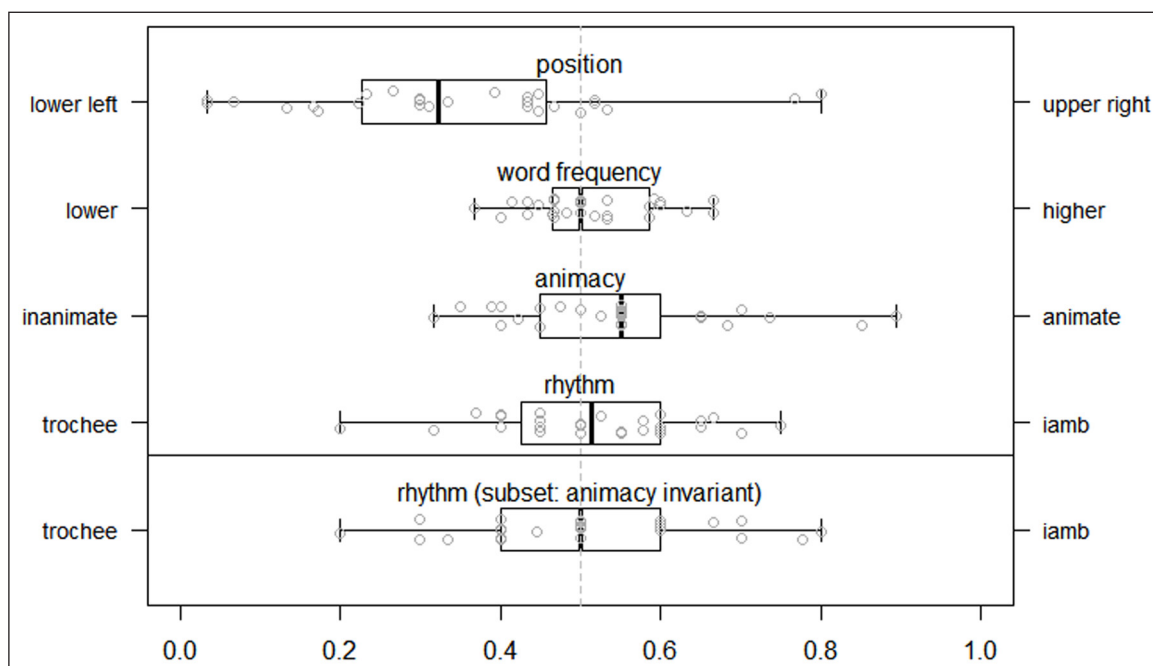


Figure 11: Mean ratio of conjunct order in Experiment 3 (adults, conjunction with *oder*, ‘or’) with respect to the following factors: position (proportion upper right first), word frequency (proportion higher frequent first), animacy (proportion animate first), and prosody (proportion iamb first). Each plot represents the influence of one factor on word order. The bold line within each box indicates the median across participants. Grey circles represent means of individual participants. Chance distribution (.5 ratio) is indicated by the vertical dotted line. The four upper plots refer to the dataset where the relevant factors vary whereas the lowest plot refers to the subset where animacy doesn’t vary in the item pairs.

order when the data is averaged (see **Figure 11**), frequency turns out to significantly affect conjunct ordering. *LAPSE and the interaction between ANIM and *LAPSE yield non-significant results. In a second step, we isolated animacy-invariant item pairs. In this subset, as shown in **Figure 11**, there was no significant effect of rhythm or frequency. The model statistics are summarized in **Table 7**.

4.3. Discussion of Experiment 3

In Experiment 3 we set out to test whether the results which we interpreted as an effect of prosody in Experiment 1 and 2 could be alternatively explained by potential artefacts of any item-specific properties. This was done by changing the instructions that were given to the participants: Instead of using the monosyllabic conjunction *und* (‘and’), participants were instructed to use the trochaic coordinative disjunction *oder* (‘or’) between the two nouns. Consequently, both sequences of iambs and trochees yielded disrhythmic structures so that none of them should be preferred over the other (if both rhythmic constraints are equally strong, they may cancel out each other). A tendency to front iambs in this design could mean that our effects in the first two experiments aren’t necessarily rhythmic in nature but potentially caused by item-specific properties of our iambic nouns or their corresponding pictures (e.g., visual salience). To sum up, in Experiment 3 we predicted no effect of rhythm but an effect of animacy, and this is exactly what we found.

4.3.1. Comparison of subset analyses of Experiments 2 and 3

While we obtained, as predicted, a significant effect of *LAPSE in the subset analysis in Experiment 2 (see above), and no significant effect of *LAPSE in the subset analysis of Experiment 3, the difference between these two effects is relatively small, as the confidence

Table 7: Results of the generalized linear mixed-effects regression model (glmer) with rhythm, animacy, their interaction, and word frequency as fixed effects. The top part shows the results for the whole data set; the bottom part shows the results for the subset of item pairs in which the two nouns are either both animate or both inanimate.

Model (glmer) for complete data set				
	Estimate	SE	z value	p value
(Intercept)	-0.72857	0.19304	-3.774	<0.001***
ANIM	0.44683	0.19626	2.277	0.022801*
*LAPSE	0.26963	0.20020	1.347	0.178050
ANIM_Diff	-0.03651	0.17466	-0.209	0.834434
Freq_Diff_log	0.50809	0.17697	2.871	0.004091**
ANIM:*LAPSE	0.72189	0.57480	1.256	0.209151
*LAPSE:ANIM_Diff	0.15071	-0.38680	-0.390	0.696811
Model (glmer) for animacy-invariant subset				
	Estimate	SE	z value	p value
(Intercept)	-0.793	0.284	-2.790	0.0055**
*LAPSE	0.365	0.298	1.227	0.219
Frequency	0.486	0.342	1.419	0.156

intervals (coefficient estimate $\pm 2xSE$) associated with these effects clearly overlap to a non-negligible amount (confidence interval for *LAPSE in subset analysis of Experiment 2: [0.086, 1.226], Experiment 3: [-.231, .961]).⁹ It is therefore possible that a re-run of Experiment 3 would yield an effect of *LAPSE as well.

Recall that we argued that any effect of stress lapse in trochee-initial phrases (*Káfer óder Klavíer*) would be neutralized by stress clash in iamb-initial phrases (*Klavíer óder Káfer*). Furthermore, the large overlap in the confidence intervals of the coefficients might well have to do with the fact that *oder*, in spite of its trochaic stress pattern, usually receives only very limited prosodic prominence on its stressed syllable and certainly no accent. The effect of stress clash in case of iamb-initial phrases might therefore be minuscule and not strong enough to override the effect of stress lapse in trochee-initial phrases.¹⁰ Therefore, we cannot exclude that iamb-initial phrases with *oder* might be felt to be

⁹ A reviewer suggested to pool the animacy-invariant data subsets of both Experiment 2 and Experiment 3 and run a model with 'Experiment' as a factor and its interaction with *LAPSE included as fixed effects. This model yields neither a significant effect of LAPSE nor a significant interaction of *LAPSE:Experiment. That is, the main effect of *LAPSE in Experiment 2 is too weak to significantly show when the data of Experiment 3 adds variability. At the same time, the difference in the effect of LAPSE between the Experiments 2 and 3 is too small for the interaction to come out significant as well.

Model (glmer) for animacy-invariant subset

	Estimate	SE	z value	p value
(Intercept)	-0.78	0.267	-2.927	0.0034**
*LAPSE	0.34	0.287	1.199	0.231
Experiment	-0.14	0.362	-0.405	0.686
Frequency	0.43	0.236	1.817	0.069
*LAPSE:Experiment	0.33	0.388	0.841	0.401

¹⁰ In the introduction, we illustrated that the rhythmic constraint *LAPSE generally implies a prosodic violation (unparsed syllables), whereas *CLASH seems to be purely rhythmical (see Figure 2). A stronger effect of *LAPSE as opposed to *CLASH corroborates this account, i.e., *LAPSE interacts with higher levels of the prosodic hierarchy, while *CLASH only affects the local sequence of (stressed and unstressed) syllables.

x					x		x			x		3
x		x			x		x	x		x		2
x	x	x	x	x	x		x	x	x	x	x	1
Kä- <i>beetle</i>	fer <i>or</i>	o- <i>or</i>	der <i>piano</i>	Kla- <i>piano</i>	vier <i>piano</i>		Kla- <i>piano</i>	vier <i>or</i>	o- <i>or</i>	der <i>beetle</i>	Kä- <i>beetle</i>	fer <i>beetle</i>

Figure 12: Metrical grid with three layers. Layer 1 represents the default beat for every syllable. Layer 2 represents the beat for every stressed syllable. Layer 3 represents the beat for every accent bearing syllable. The grid illustrates a stress lapse (left panel) and a stress clash (right panel). The (potential) violations of *LAPSE and *CLASH are highlighted in bold.

nearly as rhythmic as those with *und*. **Figure 12** illustrates the prominence relations in a metrical grid, including a third level covering the accent bearing nouns as opposed to the conjunction that bears lexical stress but no accent (see Selkirk, 1984, among others). The difference in terms of accentability might weaken the stress clash in this structure. However, the stress lapse is not affected by this.

5. General Discussion

In the present study, we examined the main effects and interaction of two constraints, ANIM and *LAPSE, on German preschoolers’ and adults’ production of conjoined bare noun phrases. The semantic constraint ANIM states that animate items should be produced before inanimate ones, whereas the rhythmic constraint *LAPSE requires the avoidance of two or more unstressed syllables in a sequence. In Experiments 1 and 2, preschoolers and adults were instructed to name pairs of pictures using conjoined bare noun phrases (e.g., ‘dolphin and planet’) without any prespecified order of the conjuncts. Due to the structure of the stimulus material, producing those noun phrases combined with a monosyllabic unstressed *und* (‘and’) resulted either in violations of *LAPSE (*Rátte und Planét*, ‘rat and planet’), or ANIM (*Planét und Rátte*, ‘planet and rat’), both (*Hóse und Delfín*, ‘trousers and dolphin’), or yielded phrases that comply with both the phonological and the semantic constraint (*Delfín und Hóse*, ‘dolphin and trousers’). In Experiment 3, the adult participants were instructed to name the same pairs of pictures again, but this time with the trochaic disjunction *oder* (‘or’) between the nouns. With the exchange of the conjunction, all phrases involving at least one iamb yielded a rhythmically suboptimal structure irrespective of their word order, i.e., a stress clash (*Klavíer óder Káfer*, ‘piano or beetle’) or a stress lapse (*Káfer óder Klavíer*, ‘beetle or piano’).

The remainder of the General Discussion starts with summarizing the results in Section 5.1. Afterwards, Section 5.2 discusses possible interactions of lower-level rhythm with higher-level prosody which might have affected the results. In Section 5.3 we integrate metrical grids and discuss possible rhythmic accommodation in our structures. In Section 5.4., we consider that trochaic and animate items might be retrieved faster than the iambic and inanimate ones, which could confound our tested constraints on word order. In Section 5.5, we discuss the different needs of forward planning of both constraints in order to influence word order. We elaborate on this in the context of speech production models in Section 5.6, specifically integrating the task of naming two isolated objects in a conjoined noun phrase in Section 5.7. A final summary in Section 5.8. concludes the paper.

5.1. The results in a nutshell

In Experiment 1 children preferably produced animate items before inanimate ones, showing a significant influence of ANIM on word order. These results are consistent with findings reported by Prat-Sala et al. (2000) for English speaking children. The effect of

ANIM correlated positively with age, which is rather surprising in preschoolers, given that Bassano et al. (2013) reported ceiling effects of animacy in children aged only 2;5 years. However, taking possible reasons for an effect of animacy into account, this result is plausible: For example, the notion of ‘concept accessibility’ (Bock & Warren, 1985) assumes the higher number of semantic pathways leading to an ‘animate’ lexical entry to accelerate word access. Since the lexical entries are getting richer and more diverse during language development, an effect of animacy due to concept accessibility might increase with age. Also, syntactic and syntactic-semantic accounts which see the effect of ANIM in relation to typical requirements of the usually sentence initial subjects or agents, are in accordance with this result—older children have more experience with these constructions which most likely increased their influence on children’s serialization.

Furthermore, participants preferred to name iambs before trochees, and by doing so, avoided disrhythmic structures. However, the rhythmic effect only reached the conventional level of significance for the subset of data in which the two referents did not vary with respect to animacy. Consequently, the experiment lends some support to our first hypothesis, in line with the *prosodic licensing hypothesis* (Demuth, 2007). That is, while conjunct order is demonstrably affected by ANIM and *LAPSE, the scope of the first constraint clearly surpasses the second one. In Experiment 2, adult participants showed a remarkably similar pattern in sequencing nouns with ANIM dominating *LAPSE. These results add to, and corroborate, the findings by McDonald et al. (1993) for English speaking adults. In Experiment 3, we did not find any effect of word stress—since both sequences yielded a disrhythmic structure, this pattern was predicted. Thus, this null result is in line with our prosodic interpretation of the results in Experiments 1 and 2. The effect of ANIM remained significant in Experiment 3. In the following, we will discuss potential weaknesses of our design, and explore the workings of the semantic and rhythmic constraints that might offer possible explanations for the relatively weak effect of *LAPSE.

5.2. Potential prosodic artefacts

As described in the introduction, this study was inspired by McDonald et al. (1993), who tested adult English participants and found that both animacy and rhythm affect the word order of conjuncts—the latter only as long as animacy doesn’t vary. Crucially, in their experiment, they used disyllabic and monosyllabic items. To exclude any effect of word length (Lohmann & Takada, 2014), we used two types of disyllabic items—trochaic and iambic ones. In the following, we will outline the potential prosodic and psycholinguistic consequences of this choice.

As described above, participants had to name iambs before trochees to produce a rhythmically alternating sequence in Experiments 1 and 2, which yielded an unstressed initial syllable. Note, however, that even the structure that yields rhythmic alternation involves a prosodic violation. This is illustrated in **Figure 13**: In the rhythmically alternating structure the first syllable *Pi-* cannot be parsed into a foot, leading to a violation of exhaustivity. Since Gerken (1994a, among others) showed, that especially children avoid (initial) unparsed syllables, our design might provoke counteracting effects of prosody that could explain that the expected effect of rhythm remained rather weak.

At this point, it is worthwhile to stress that despite the issue of an unparsed initial syllable, we actually found a rhythmic effect. We might conclude that the preference for rhythmic alternation had a stronger influence on the chosen word order than the violation evoked by an unparsed initial syllable. Additionally, the prosodic structure in the case of rhythmic alternation (**Figure 13**, left panel) violates the exhaustivity constraint only once while the sequence with a lapse does so twice (**Figure 13**, right panel): In the lapse structure, the syllables *Pi-* and *und* can’t be parsed into a foot yielding two violations of

exhaustivity. Gerken (1996) showed that the probability to omit unstressed syllables, correlates with the number of exhaustivity-violations in young children. Thus, while not prosodically perfect, the rhythmically alternating structure is prosodically more well-formed than the disrhythmic structure. This may have led to a (small) preference for the former compared to the latter.

5.3. Potential rhythmic process

This study is built on the assumption that the German conjunction *und*, ‘and’ is unstressed. However, in a sentence repetition task, Vogel, van de Vijver, Kotz, Kutscher, and Wagner (2015) showed that the prominence degree of usually unstressed function words varies in favor of an alternating rhythm. The authors found increased syllable durations of the pronoun *es*, ‘it,’ if surrounded by unstressed syllables. Transferring their finding to our target structure, the originally unstressed conjunction in our disrhythmic sequences (surrounded by two unstressed syllables) might be processed similarly to the originally unstressed pronoun in Vogel et al. (2015)—that is, with a higher level of prominence. **Figure 14** illustrates this possibility in a metrical grid (for rhythmic adjustments in metrical grids see Nespor & Vogel, 1986). However, the fact that we still found an effect of *LAPSE suggests that any possible phonetic adjustment (beat insertion) did not invalidate our assumptions.

5.4. Potential artefacts of word form retrieval

Another potentially confounding factor is related to the costs of word form retrieval. Given that, in German, the iambic stress pattern is less frequent than the trochaic one, iambic words may result in increased processing effort in speech production. Since, in

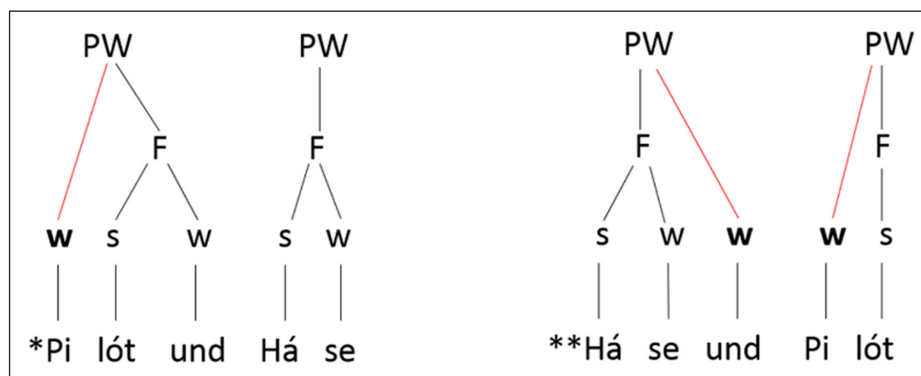


Figure 13: Violations of exhaustivity in the prosodic hierarchy for both word orders in Experiment 1 and 2. The red lines represent ‘illegal’ syllable parsing. The left panel shows a rhythmically alternating sequence, the right panel a *LAPSE sequence.

x				x					x	3
x				x			x		x	2
x	x	x	x	x		x	x	x	x	1
Rat-	te	und	Re-	gal	Rat-	<i>te</i>	<i>und</i>	<i>Re-</i>	gal	
‘rat		and	shelf’		‘rat		and	shelf’		

Figure 14: Metrical grid with three layers. Layer 1 represents the default beat for every syllable. Layer 2 represents the beat for every stressed syllable. Layer 3 represents the beat for every accented syllable. The left panel illustrates the metrical grid of the structure violating *LAPSE (in bold). The right panel illustrates the sequence with a rhythmically motivated beat insertion on the syllable *und*, ‘and’ resulting in a rhythmically alternating structure.

our design, participants had to name iambs before trochees to produce a rhythmically optimal phrase (*Klavier und Rátte*, ‘piano and rat’), increased processing difficulty with the iambic structure could explain the rather weak effect of *LAPSE on serialization. Indeed, Schiller et al. (2004) found that English speaking adults need more time for naming iambs than they do for naming otherwise comparable trochees. Consequently, there could be counteracting effects of prosody and word form retrieval with shorter naming latencies for trochees (trochee > iamb) and a preference for naming iambs first to yield rhythmic structures (iamb > trochee).

Confirming the particular difficulty with iambic words (see also Gerken, 1994a) the children in Experiment 1 made more mistakes with iambic items than with trochees, although both were carefully familiarized in the same way. For example, participants made some mistakes naming the pictures for *Ballón* (‘balloon’), *Regál* (‘shelf’), *Gespénst* (‘ghost’), and *Pilót* (‘pilot’) using synonyms like *Géist* for *Gespénst* (‘ghost’), or semantically related words like *Schránk* (‘cupboard’) for *Regál* (‘shelf’), or *Flúgzeug* (‘plane’) for *Pilót* (‘pilot’). In some cases, they used more specific words like *Héiþluftballon* (‘hot-air balloon’) for *Ballón* (‘balloon’) or they couldn’t find any word to name the picture at all. A phrase with any of these variations was considered as invalid yielding less data points in structures involving iambs.

In any case, the effect of rhythm was detectable despite this possible counteracting effect of word form retrieval. Nevertheless, one might argue that at least in Experiment 1, which was conducted with children, the effect of rhythm might have been stronger without the differing degree of effort retrieving trochees and iambs.¹¹

5.5. Semantic and rhythmic constraints and the scope of speech planning

The effect of ANIM in our experiment can probably be attributed to the lexical-semantic level and possibly also to visuo-semantic processing of the words and pictures involved. When producing isolated conjoined noun phrases, participants did not have to assign the items to any thematic roles or syntactic functions. Therefore, explanations relating to the syntactic level seem unlikely. Rather, it seems plausible to assume an easier lexical access for animate compared to inanimate items. Indeed, for adult participants, Proverbio, Del Zotto, and Zani (2007) found that naming latencies for animate words are faster than for inanimate ones. Such an eased access could be due to a larger quantity of semantic features or pathways leading to these features for animate as opposed to inanimate items (Bock & Warren, 1985). More specifically, an eased access could be related to increasing grades of a referent’s animacy (Yamamoto, 1999) enhancing its ‘concept accessibility’ (Bock & Warren, 1985). For example, given the criterion of self-paced movement (‘self-propelledness’), it seems to be reasonable that the attributed ability of autonomous movement may facilitate access to the lexical entry of a referent, as it can be associated with different situations and in different motions and has therefore more semantic pathways leading to it compared to an inanimate object.

Furthermore, effects of ANIM on word order in naming two pictures in a conjoined noun phrase could be due to visual salience. In an eye tracking study, Carniglia, Carputi, Manfredi, Zambarbieri, and Pessa (2012) showed that there is a bias for fixations to pictures with animate referents compared to those with inanimate referents. The former were earlier, longer, and more often fixated than the latter. As a consequence, our effect

¹¹ Surprisingly, Gladfelter and Goffmann (2013) found that English speaking preschoolers are more successful in learning new iambs than in learning new trochees. The authors found more precision in pronouncing new iambic words as well as a higher score in naming them correctly. They assume that—for trochees—children rely on earlier learned and hence, potentially less precise motor plans. The fact that the items in the present study weren’t ‘new’ for the children may have contributed to their ‘trochaic bias.’

of ANIM might also be due to a visual bias in favor of animate items, with a possibility of a *first seen – first planned – first named* – mechanism. With visual prominence being controlled in our material in a pretest, an explanation of ANIM just based on visual complexity or dominance is not very likely though.

In sum, it appears that an attentional bias for animate objects (that may be grounded in some kind of visual and/or semantic salience) would suffice to explain the animacy effect in our data. Accordingly, speakers would name the more attention-grabbing animate noun first and then move on to name the second object. It would thus not be necessary for participants to directly compare the semantic features of the two objects. In order to produce rhythmically alternating structures on the other hand, speakers have to consider and compare the prosodic makeup of both nouns and the whole phrase. To abide by the rhythmic constraint, the scope of speech planning therefore needs to be considerably larger as opposed to the animacy constraint (for further discussion in the context of speech production models, see Section 5.6).

However, other explanations cannot be excluded. In particular, a tendency to produce animate nouns before inanimate ones could also be based on statistical properties of the speech input. This input typically consists of sentences and phrases so that syntactic or syntactic-semantic mechanisms, like the usual order of theta roles or syntactic functions, could be indirectly responsible for tendencies in the serialization of conjoined nouns ('indirect syntactic explanation'). Further research is needed to clarify this question.

5.6. Semantic and rhythmic constraints in a speech production model

Apparently, ANIM and LAPSE are effective on different levels of speech planning. Which levels these are and how our results could fit into a model of speech production will be discussed in the following. First, we need to address the fact that our rather artificial structure of two bare noun phrases neither fits into a model of single word production, nor does it need all aspects of a model for whole sentences or beyond.

5.6.1. The model by Bock and Levelt (1994)

Classical speech production models (e.g., Bock & Levelt, 1994) postulate sentence planning to proceed incrementally on different levels starting with a preverbal message. This preverbal message is linguistically encoded during grammatical and phonological encoding. The former is further divided into functional processing, where lemmas are selected with their syntactic functions in this specific sentence and positional processing which generates a first syntactic structure with the linear order of the constituents and their inflection. The latter integrates the retrieval of the wordforms with their segments, metrical frame, and also the prosodic structure of the sentence (Levelt, 1989). Noteworthy, these models work on the assumption that each level "is influenced only by information represented at the level directly above it." (Bock & Levelt, 1994, p. 949) Consequently, feedback from phonological to grammatical encoding is not envisaged.

Transferring our structure into such a model, we obviously don't need to account for function assignment during functional processing, as syntactic functions are not specified in conjunctive phrases, nor do we need a process for inflection in positional encoding. However, what we do need from grammatical encoding, is lemma selection from the visual stimulus to account for lexical access of the nouns. Crucially, we also need constituent assembly to account for the serial order of the nouns. From phonological encoding, wordform access and presumably the generation of a rhythmic and prosodic structure of the conjunctive phrase need to be integrated in the simplified model.

Finally, we need to specify the level at which the constraints, ANIM and *LAPSE, are effective—and how it is possible that both constraints have at least some impact

on sequencing the nouns (which is assumed to be fixed during positional processing in grammatical encoding). Following Bock and Levelt (1994), animacy could be located as a preverbal feature of a message (together with, among others, the visual form). Consequently, the impact of ANIM would precede positional processing, which makes its influence on ordering the nouns easily applicable. Alternatively, assuming animacy to be part of the semantic information stored in the lexical entry, its impact on serial order is still consistent with the model, since lemma access also precedes positional encoding. Thus, the influence of ANIM can be easily accounted for in a speech production model. The integration of the rhythmic constraint is much more challenging, as it showed some impact on sequencing the nouns, at least when animacy was ruled out as a varying factor.

In a model as proposed by Bock and Levelt (1994), *LAPSE is not predicted to influence word order, since rhythm and prosody are only accounted for during phonological encoding which is predicted to happen only after positional processing. (See also the prosody generator in Levelt, 1989). One way to address this issue, is to integrate a monitoring loop into the model, as proposed by Levelt (1983). However, since our participants were instructed to name the pictures in a swift way, such a loop doesn't seem very likely to influence word order. As a consequence, our effect of *LAPSE seems unexplained by the model in its discussed form. Alternatively, we could assume that the model does not work serially, and that a processing level can be influenced not only by information represented at the level above it (feed forward), but also by information represented at the level below it (feedback), i.e., if we assume interactive processing. In this way, rhythm could influence positional processing. However, given that the model proposed by Bock and Levelt (1994) is strictly serial, no feedback should occur.

5.6.2. The model by Keating and Shattuck-Hufnagel (2002)

Another solution was proposed by Keating and Shattuck-Hufnagel (2002) who integrate basic functions from the models introduced above while taking a 'prosody-first-view.' Their key assumption is that prosodic encoding precedes phonological encoding and thus has more influence on structural aspects of an utterance than it does in the classical view. However, the authors do not explicitly mention the change of word order as a form of restructuring. Rather, they refer, for example, to boundary tones, phrasal accents, additional boundaries, and cliticization. In the following, we will sketch where a reordering of words due to rhythmic constraints might be located at, following their model.

According to the authors, an initial default prosodic structure is derived from the syntactic surface structure—at this point, still not containing any word form information. This default prosodic structure then underlies a cyclic process of restructuring, a process which is roughly following the domains of the prosodic hierarchy, starting from the highest level with intonational phrases (IP) followed by phonological phrases (PP) and prosodic words (PW). At each level, the word form information which is required for restructuring, is becoming richer and more detailed.

During IP formation the authors assume boundary tones to be computed as well as the number of words (still without their metrical frames). During PP formation, the default phrasal accent and the metrical frames of each lemma (being most interesting for our rhythmic constraint, *LAPSE) are computed or retrieved. Consequently, restructuring based on word stress and rhythm should be located at this specific cycle. However, information about foot structure is not yet available at this stage, so that low-level rhythm like *LAPSE avoidance probably does not influence the structure of the phrase, i.e., word order which is relevant for this study. Furthermore, we do not think that rhythm is predicted to influence word order at this stage because not all syllables are supposed to be retrieved. This makes rhythm-based constraints ineffective at this level: "In general we do not have final stress patterns at this point: because we haven't conjoined the metrical

frames of the various morphemes, we haven't made any stress adjustments yet." (Keating & Shattuck-Hufnagel 2002, p. 145).

PW formation involves processes like cliticization (which might be relevant for the unstressed conjunction *und*, 'and,' in our design: *Löwe und Pilot*, 'lion and pilot'). Also, the corresponding syllables of affixes are supposed to be computed here. Consequently, syllables with their stress value are supposed to be retrieved so that changes of word order due to the current rhythmical structure might occur at this stage. However, changes of word order motivated by prosody are not mentioned, let alone changes of word order due to lower-level rhythmic constraints. Therefore, an explanation of the rhythmic effect observed in the present experiments remains unclear. Further, the authors explicitly locate rhythmic adjustments at the level of phonological encoding where changes of word order are not predicted.

To sum up, we also cannot explain a rhythmic effect on word order in the context of this model. In the following, we discuss a further model which was largely inspired by Keating and Shattuck-Hufnagel (2002), although taking a different angle with a larger focus on Bock and Levelt (1994).

5.6.3. The model by Calhoun (2010)

Calhoun (2010) adopted Bock and Levelt's model and inserted some additional components while keeping their basic assumption of a unidirectional transmission of information between grammatical and phonological encoding. One of their main changes is a component in grammatical encoding for the generation of 'high level prosodic structure.' This includes breaks between IPs as well as nuclear accents and interacts with the two other components, one of them being positional processing. According to this account, the output of grammatical encoding is a surface structure which is, similar to Keating and Shattuck-Hufnagel (2002), prosodic in nature (including nuclear accents). At the level of phonological encoding, she inserted a prosody regulator (replacing Levelt's prosody generator). This component integrates the metrical information of lexemes into the higher-level prosodic structure (received via the surface structure) and operates on low level constraints like rhythmic alternation. Further, Calhoun (2010) keeps Levelt's metrical spellout which retrieves the metrical structure of words in the first place (and integrates the expected prominence of a word, given its lexical class).

In the present case, an isolated conjunctive phrase like *Löwe und Pilot* usually corresponds to a single IP. In the modified model by Calhoun (2010), IP-specific prosodic information is represented already during grammatical encoding, including major breaks and nuclear accents. Also, lemma entries are retrieved here, so that these can be associated with a nuclear accent (if positional processing locates the lemma node at the right edge of the IP). In the structure of the present study, this means that in a conjunctive phrase like *Löwe und Pilot*, 'lion and pilot,' *Pilot* should be marked with a nuclear accent. However, since wordform retrieval and rhythmic constraints are not supposed to be working until phonological encoding, more specifically, until the prosody regulator and the metrical spellout retrieve the surface structure with the corresponding lemma nodes, the nuclear accent is just associated with the word, not with a specific syllable. Thus, the placement of the nuclear accent as well as positional encoding cannot be determined by the metrical structure of the word. As a consequence, an impact of *LAPSE on word order is still not expected, and thus, we cannot explain its effectiveness in the discussed serial models of speech production.¹²

¹² Note that the model by Calhoun (2010) was not intended to rule out word-level metrical effects on word ordering. Rather, the proposal was a sketch and issues related to *LAPSE were not relevant to the main topic in that paper, so this was not really considered (Calhoun, personal conversation).

One possible solution to this problem might be to assume that our instruction of naming the pictures in a swift way did not hinder silent monitoring and sequence rearrangement. Further, an interactive model would allow for interactions between higher-level prosodic encoding and lower-level rhythmic optimization (Dell & O’Seaghdha, 1992). Alternatively, modifying the model by Calhoun (2010), the prosodic component in grammatical encoding might include not only nuclear accents, but also the metrical structure of the words (as it does in Keating & Shattuck-Hufnagel, 2002), and with these, the implementation of rhythmic constraints. Although this idea is somewhat unconventional in the context of serial models, it may be possible for the specific task of naming two ‘independent’ pictures. In the following section we will address this issue.

5.7. *The present task in a speech production model*

Apart from the artificial character of isolated conjoined noun phrases, the task of naming two objects in response to a visual stimulus differs from encoding some intention or coherent preverbal message: Participants have to name two unconnected objects in response to two drawings. This means they need to retrieve two semantically and syntactically unrelated lexical entries. Meyer, Sliderink, and Levelt (1998, referred to as MSL) conducted an eye tracking study with the exact target structure (two nouns without determiner with ‘and’ in between). Their stimuli were also two drawings of objects but placed next to each other (left and right) on the screen. They found that “the viewing times depended systematically on the time needed to retrieve the phonological form of the object names.” (MSL, p. 32).

According to their results, “the speakers completed the conceptual and most of the linguistic processing for one object before initiating the shift of gaze to the next object” (MSL, p. 32). Additionally, about 50% of their participants moved their gaze from the second picture back to the first picture during speech onset (MSL, p. 32). This implies, following the authors, that those participants had retrieved also the second word form by the time they started speaking.

Transferring these findings to the example *Löwe und Pilot* (‘lion and pilot’), the speaker has already retrieved the complete lexical entry of the first fixated picture *Löwe* (‘lion’) before she starts to look at the second one, *Pilot* (‘pilot’). Also, she has retrieved some parts of the lexical entry of *Pilot* before she starts speaking. Assuming that *Löwe und Pilot* (or *Pilot und Löwe*) is an IP which is computed as a single unit during grammatical encoding (following Calhoun, 2010), the speaker already retrieved the phonological form of one word and some information about the lexical entry of the other word while computing this IP. Consequently, she already has access to the metrical structure of one word and possibly of the other. In those cases, in which she looked back at the first picture, she should have accessed the phonological form of both words (according to MSL, gaze shifts back to the first item were only observed after full word access of the second item). If the latter is the case, rhythmic constraints could be effective already during IP formation—which interacts with positional processing, and thus with word order.

In the present example, the disrhythmic sequence *Löwe und Pilot* could be ‘rearranged’ into the rhythmically alternating structure *Pilot und Löwe* during grammatical encoding where prosody interacts with positional processing. Due to the specific task, it is possible that some participants already retrieved the metrical structure of both words when IPs are constructed, allowing rhythmic constraints to be effective during grammatical encoding. Consequently, in the context of this specific task and target structure, the effect of *LAPSE could be integrated into the model. In this framework it also makes sense that ANIM shows a relatively strong effect (presumably already operative in the preverbal message),

while the effect of *LAPSE is much weaker—only happening during IP formation in those cases in which the metrical frame of both words has been retrieved already.

5.8. Final summary

This study demonstrates that both animacy and rhythm affect the ordering of bare noun conjuncts in German, and they do so in both preschoolers' and adults' speech production. We found a consistent effect of animacy (a preference for naming animate referents before inanimate ones) in all experiments, in line with results by Prat-Sala et al. (2000, among others). Effects of rhythm (i.e., the preference for conjunct orders that yield a rhythmic alternation of stressed and unstressed syllables) were only detectable when the animacy of the conjuncts involved was held constant. This result is consistent with findings by McDonald et al. (1993). Our experiments show that both the animacy effect and the rhythmic effect hold for preschoolers as well as adult speakers of German. For the group of preschoolers, the preference for naming animate referents first increases with age. We have no evidence for a corresponding developmental change regarding the influence of rhythm on word order choice.

Additional Files

The additional files for this article can be found as follows:

- **Appendix.** PDF file containing the experimental lists and scatterplots. DOI: <https://doi.org/10.5334/labphon.254.s1>
- **Supplementary Material.** ZIP file containing the data and analysis code. DOI: <https://doi.org/10.5334/labphon.254.s2>

Acknowledgements

This research is part of the doctoral dissertation to be submitted by the first author. We thank Markus Bader for his invaluable help with the planning of the Experiments 2 and 3 as well as their interpretation. Karen Henrich, Yvonne Portele, Vasiliki Koukoulioti, Alice Schäfer, Beata Moskal, Caroline Féry, Julia Biskupek, Frank Kügler, Stefan Blohm, and Christine Knoop provided helpful feedback at various stages of the project. Marc Schwab, Liane Pietsch, and Anna Pressler helped coding the audio data. Debby Trzeciak helped with the pictures shown in this paper. We are also grateful for constructive comments of three anonymous reviewers. We thank the editors for putting together this special issue and inviting us to submit the manuscript and improving it substantially with their comments. Last but not least, we thank all adult—and child participants, their parents, and the kindergartens for their patient cooperation.

Funding Information

Parts of this research were funded through a grant of the German Research Foundation to Gerrit Kentner (DFG grant no. KE 1985/2–1).

Competing Interests

The authors have no competing interests to declare.

References

- Anttila, A. (2016). Phonological effects on syntactic variation. *Annual Review of Linguistics*, 2, 115–137. DOI: <https://doi.org/10.1146/annurev-linguistics-011415-040845>
- Bassano, D., Korecky-Kröll, K., Maillochon, I., Van Dijk, M., Laaha, S., Van Geert, P., & Dressler, W. U. (2013). Prosody and animacy in the development of noun determiner

- use: A cross-linguistic approach. *First Language*, 33(5), 476–503. DOI: <https://doi.org/10.1177/0142723713503253>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using {lme4}. *Journal of Statistical Software*, 67(1), 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>
- Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21, 47–67. DOI: [https://doi.org/10.1016/0010-0277\(85\)90023-X](https://doi.org/10.1016/0010-0277(85)90023-X)
- Bock, K., & Levelt, W. J. (1994). *Language production: Grammatical encoding* (pp. 945–984). Academic Press.
- Bohn, K., Knaus, J., Wiese, R., & Domahs, U. (2013). The influence of rhythmic (ir) regularities on speech processing: Evidence from an ERP study on German phrases. *Neuropsychologica*, 51(4), 760–771. DOI: <https://doi.org/10.1016/j.neuropsychologia.2013.01.006>
- Boyle, M. K., & Gerken, L. (1997). The Influence of Lexical Familiarity on Children's Function Morpheme Omissions: A Nonmetrical Effect? *Journal of Memory and Language*, 36(1), 117–128. DOI: <https://doi.org/10.1006/jmla.1996.2478>
- Branigan, H. P., Pickering, M. J., & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua*, 118(2), 172–189. DOI: <https://doi.org/10.1016/j.lingua.2007.02.003>
- Breiss, C., & Hayes, B. (2020). Phonological markedness effects in sentence formation. *Language*, 96(2), 338–370. DOI: <https://doi.org/10.1353/lan.2020.0023>
- Büring, D. (2013). Syntax, information structure and prosody. *The Cambridge handbook of generative syntax*, 860–895. DOI: <https://doi.org/10.1017/CBO9780511804571.029>
- Byrne, B., & Davidson, E. (1985). On putting the horse before the cart. Exploring conceptual bases of word order via acquisition of a miniature artificial language. *Journal of Memory and Language*, 24(4), 377–389. DOI: [https://doi.org/10.1016/0749-596X\(85\)90035-X](https://doi.org/10.1016/0749-596X(85)90035-X)
- Calhoun, S. (2010). How does informativeness affect prosodic prominence? *Language and Cognitive Processes*, 25(7–9), 1099–1140. DOI: <https://doi.org/10.1080/01690965.2010.491682>
- Carniglia, E., Carputi, M., Manfredi, V., Zambarbieri, D., & Pessa, E. (2012). The influence of emotional picture thematic content on exploratory eye movements. *Journal of Eye Movement Research*, 5(4), 1–9. DOI: <https://doi.org/10.16910/jemr.5.4.4>
- Clarke, A. D. F., Elsner, M., & Rohde, H. (2015). Giving Good Directions: Order of Mention Reflects Visual Salience. *Frontiers in Psychology*, 6, 1–10. DOI: <https://doi.org/10.3389/fpsyg.2015.01793>
- Cavalho, A., Dautriche, I., & Christophe, A. (2015). Preschoolers use phrasal prosody online to constrain syntactic analysis. *Developmental Science*, 19(2), 235–250. DOI: <https://doi.org/10.1111/desc.12300>
- Cavalho, A., Lidz, J., Tieu, L., Blears, T., & Christophe, A. (2016). English-speaking preschoolers can use phrasal prosody for syntactic parsing. *The Journal of the Acoustical Society of America*, 139(6). DOI: <https://doi.org/10.1121/1.4954385>
- Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, 42(1–3), 287–314. DOI: [https://doi.org/10.1016/0010-0277\(92\)90046-K](https://doi.org/10.1016/0010-0277(92)90046-K)
- Demuth, K. (2007). Acquisition at the Prosody-Morphology Interface. In: A. Belokova, L. Meroni, & M. Umeda (Eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, 84–91. Somerville, MA: Cascadia Proceedings Project. www.lingref.com, document #1549.

- Demuth, K., Machobane, M., Moloi, F., & Odato, C. (2005). Learning Animacy Hierarchy Effects in Sesotho Double Object Applicatives. *Language*, *81*, 421–427. DOI: <https://doi.org/10.1353/lan.2005.0056>
- DeRuiter, L., Narasimhan, B., Chen, J., & Lack, J. (2018). Children's use of prosody and word order to indicate information status in English noun phrase conjuncts. *Proceedings of the Linguistic Society of America*, *3*. DOI: <https://doi.org/10.3765/plsa.v3i1.4331>
- Dewart, M. H. (1979). The role of animate and inanimate nouns in determining sentence voice. *British journal of psychology*, *70*, 135–141. DOI: <https://doi.org/10.1111/j.2044-8295.1979.tb02151.x>
- Domahs, F., Blessing, K., Kauschke, C., & Domahs, U. (2016). Bono Bo and Fla Mingo: Reflections of speech prosody in German second graders' writing to dictation. *Frontiers in Psychology*, *7*, 856. DOI: <https://doi.org/10.3389/fpsyg.2016.00856>
- Drenhaus, R., & Féry, C. (2008). Animacy and child grammar: An OT account. *Lingua*, *118*, 222–244. DOI: <https://doi.org/10.1016/j.lingua.2007.02.006>
- Eimas, P. D. (1996). The Perception and Representation of Speech by Infants. In J. L. Morgan & K. Demuth (Eds.). *Signal to Syntax*. Lawrence Erlbaum.
- Fox, A. V. (2006). TROG-D. Test zur Überprüfung des Grammatikverständnisses. Handbuch. Das Gesundheitsforum Idstein: Schulz-Kirchner Verlag.
- Fox-Boyer, A. V. (2015). *Kindliche Aussprachestörungen: Kindliche Aussprachestörungen*. Schulz-Kirchner Verlag GmbH.
- Gámez, P. B., & Vasilyeva, M. (2015). Exploring interactions between semantic and syntactic processes: The role of animacy in syntactic priming. *Journal of Experimental Child Psychology*, *138*, 15–30. DOI: <https://doi.org/10.1016/j.jecp.2015.04.009>
- Gelman, S. A., & Gottfried, G. (1996). Children's causal explanations of animate and inanimate motion. *Child Development*, *67*, 1970–1987. DOI: <https://doi.org/10.2307/1131604>
- Gerken, L. A. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language*, *30*(4), 565–584. DOI: [https://doi.org/10.1016/0749-596X\(91\)90015-C](https://doi.org/10.1016/0749-596X(91)90015-C)
- Gerken, L. A. (1994a). A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language*, *21*(3), 565–584. DOI: <https://doi.org/10.1017/S0305000900009466>
- Gerken, L. A. (1994b). Young children's representation of prosodic structure: Evidence from English-speakers' weak syllable omissions. *Journal of Memory and Language*, *33*, 19–38. DOI: <https://doi.org/10.1006/jmla.1994.1002>
- Gerken, L. A. (1996). Prosodic structure in young children's language production. *Language*, *72*, 683–712. DOI: <https://doi.org/10.2307/416099>
- Grewe, T. (2007). The Neuronal Reality of the Nominal Hierarchy: fMRI Observations on Animacy in Sentence Comprehension. Doctoral dissertation, Philipps-Universität Marburg.
- Gladfelter, I., & Goffman, L. (2013). The Influence of Prosodic Stress Patterns and Semantic Depth on Novel Word Learning in Typically Developing Children. *Language Learning and Development*, *9*(2), 151–174. DOI: <https://doi.org/10.1080/15475441.2012.684574>
- Gutman, A., Dautriche, I., Crabbé, B., & Christophe, A. (2015). Bootstrapping the Syntactic Bootstrapper: Probabilistic Labeling of Prosodic Phrases. *Language Acquisition*, *22*(3), 285–309. DOI: <https://doi.org/10.1080/10489223.2014.971956>
- Gwinner, A., Gaglia, S., & Grijzenhout, J. (2012). The effect of prosody on the acquisition of morphemes: An experimental study with German, Italian and German-Italian

- children. In: M. Grazia Busà & A. Stella (Eds.). *Methodological Perspectives in Second Language Prosody: papers from ML2P 2012* (pp. 37–41). Padova: CLEUP.
- Halle, M., & Vergnaud, J. R. (1987). Stress and the cycle. *Linguistic inquiry*, 18(1), 45–84.
- Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago, London: The University of Chicago Press.
- Keating, P., & Shattuck-Hufnagel, S. (2002). A prosodic view of word form encoding for speech production. *UCLA Working Papers in Phonetics*, 101, 112–156.
- Kentner, G., & Franz, I. (2019). No evidence for prosodic effects on the syntactic encoding of complement clauses in German. *Glossa: A journal of general linguistics*, 4(1), 1–29. DOI: <https://doi.org/10.5334/gjgl.565>
- Knudson, B., Fischer, M. H., & Aschersleben, G. (2014). Development of spatial preferences for counting and picture naming. *Psychological Research*, 79, 939–949. DOI: <https://doi.org/10.1007/s00426-014-0623-z>
- Lee, M.-W., & Gibbons, J. (2007). Rhythmic alternation and the optional complementizer in English: New evidence of phonological influence on grammatical encoding. *Cognition*, 105(2), 446–456. DOI: <https://doi.org/10.1016/j.cognition.2006.09.013>
- Legerstee, M., & Markova, G. (2008). Variations in 10-month-old infant imitation of people and things. *Infant Behavior and Development*, 31, 81–91. DOI: <https://doi.org/10.1016/j.infbeh.2007.07.006>
- Legerstee, M., Pomerleau, A., Malcuit, G., & Feider, H. (1987). The development of infants' responses to people and a doll: Implications for research in communication. *Infant behavior and development*, 10(1), 81–95. DOI: [https://doi.org/10.1016/0163-6383\(87\)90008-7](https://doi.org/10.1016/0163-6383(87)90008-7)
- Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition*, 14(1), 41–104. DOI: [https://doi.org/10.1016/0010-0277\(83\)90026-4](https://doi.org/10.1016/0010-0277(83)90026-4)
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Lohmann, A. (2013). *Constituent order in coordinate construction – a processing perspective*. Doctoral dissertation. University of Hamburg. DOI: <https://doi.org/10.1017/CBO9781139644273>
- Lohmann, A., & Takada, T. (2014). Order in NP conjuncts in spoken English and Japanese. *Lingua*, 152, 48–64. DOI: <https://doi.org/10.1016/j.lingua.2014.09.011>
- Männel, C., & Friederici, A. D. (2016). Neural correlates of prosodic boundary perception in German preschoolers: If pause is present, pitch can go. *Brain Research*, 1632, 27–33. DOI: <https://doi.org/10.1016/j.brainres.2015.12.009>
- McDonald, J. L., Bock, K., & Kelly, M. H. (1993). Word and World order: Semantic, Phonological and Metrical Determinants of Serial Position. *Cognitive Psychology*, 25, 188–230. DOI: <https://doi.org/10.1006/cogp.1993.1005>
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoni, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143–178. DOI: [https://doi.org/10.1016/0010-0277\(88\)90035-2](https://doi.org/10.1016/0010-0277(88)90035-2)
- Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, 48, 131–147. DOI: [https://doi.org/10.1016/S0749-596X\(02\)00509-0](https://doi.org/10.1016/S0749-596X(02)00509-0)
- Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, 66(2), B25–B33. DOI: [https://doi.org/10.1016/S0010-0277\(98\)00009-2](https://doi.org/10.1016/S0010-0277(98)00009-2)
- Morgan, J. L. (1986). *The MIT Press series in learning, development, and conceptual change. From simple input to complex grammar*. The MIT Press.


- Morgan, J. L., & Demuth, K. (1996). Signal to syntax: An overview. *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, 1–22.
- Narasimhan, B., & Dimroth, C. (2008). Word order and information status in child language. *Cognition*, 107(1), 317–329. DOI: <https://doi.org/10.1016/j.cognition.2007.07.010>
- Narasimhan, B., & Dimroth, C. (2017). The influence of discourse context on children's ordering of "new" and "old" information. *Linguistics Vanguard*, 4(1), DOI: <https://doi.org/10.1515/lingvan-2017-0026>
- Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, 41, 233–243. DOI: [https://doi.org/10.1016/S0167-6393\(02\)00106-1](https://doi.org/10.1016/S0167-6393(02)00106-1)
- Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht, Cinnaminson: Foris Publications.
- Opfer, J. E., & Gelman, S. A. (2011). Development of the Animate-Inanimate Distinction. Usha Gishwani (ed.). *The Wiley-Blackwell Handbook of Childhood Cognitive Development*, 213–238. Cambridge: Blackwell Publishers Ltd. DOI: <https://doi.org/10.1002/9781444325485.ch8>
- Paczynsky, M., & Kuperberg, G. R. (2011). Electrophysiological evidence for use of the animacy hierarchy, but not thematic role assignment, during verb-argument processing. *Language and cognitive processes*, 26, 1402–1456. DOI: <https://doi.org/10.1080/01690965.2011.580143>
- Piaget, J. (1978). *Das Weltbild des Kindes*. Stuttgart: Klett-Cotta.
- Pinker, S. (1984). *Language learnability and language development*. Cambridge: Harvard University Press.
- Poulin-Dubois, D., Lepage, A., & Ferland, A. (1996). Infants' Concept of Animacy. *Cognitive Development*, 11, 19–36. DOI: [https://doi.org/10.1016/S0885-2014\(96\)90026-X](https://doi.org/10.1016/S0885-2014(96)90026-X)
- Prat-Sala, M., & Branigan, H. P. (2000). Discourse Constraints on Syntactic Processing in Language Production: A Cross-linguistic Study in English and Spanish. *Journal of Memory and Language*, 42, 168–182. DOI: <https://doi.org/10.1006/jmla.1999.2668>
- Prat-Sala, M., Shillcock, R., & Sorace, A. (2000). Animacy Effects on the Production of Object-dislocated Descriptions by Catalan-speaking Children. *Journal of Child Language*, 27, 97–117. DOI: <https://doi.org/10.1017/S0305000999004031>
- Proverbio, A. M., Del Zotto, M., & Zani, A. (2007). The emergence of semantic categorization in early visual processing: ERP indexes of animal vs. artifact recognition. *BMC Neuroscience*, 8, 1–24. DOI: <https://doi.org/10.1186/1471-2202-8-24>
- R Core Team. (2020). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. URL <https://www.R-project.org/>
- Sauermann, A., & Höhle, B. (2018). Word order in German child language and child-directed speech: A corpus analysis on the ordering of double objects in the German middlefield. *Glossa: A Journal of General Linguistics*, 3(1), 57. DOI: <https://doi.org/10.5334/gjgl.281>
- Schiller, N. O., Fikkert, P., & Levelt, C. C. (2004). Stress priming in picture naming: An SOA study. *Brain and Language*, 90, 231–240. DOI: [https://doi.org/10.1016/S0093-934X\(03\)00436-X](https://doi.org/10.1016/S0093-934X(03)00436-X)
- Schröder, C., & Höhle, B. (2011). Prosodische Wahrnehmung im frühen Spracherwerb. *Sprache Stimme Gehör*, 35, 91–98. DOI: <https://doi.org/10.1055/s-0031-1284404>
- Schroeder, S., Würzner, K., Heister, J., Geyken, A., & Kliegel, R. (2014). Childlex – Eine lexikalische Datenbank zur Schriftsprache für Kinder im Deutschen. Manuskript. *MPI for human development*. Berlin. DOI: <https://doi.org/10.1026/0033-3042/a000275>

- Selkirk, E. O. (1984). Phonology and syntax: The relation between sound and structure. *Current studies in linguistics series, 10*. Cambridge, MA: MIT Press.
- Shih, S. S. (2014). Towards optimal rhythm. Doctoral dissertation. Stanford University.
- Vogel, R., van de Vijver, R., Kotz, S., Kutscher, A., & Wagner, P. (2015). 10 Function words in rhythmic optimisation. *Rhythm in cognition and grammar: A Germanic perspective, 286*, 255. DOI: <https://doi.org/10.1515/9783110378092.255>
- Wagner, M. (2010). Prosody and recursion in coordinate structures and beyond. *Natural Language & Linguistic Theory, 28*(1), 183–237. DOI: <https://doi.org/10.1007/s11049-009-9086-0>
- Wijnen, F., Krikhaar, E., & den Os, E. (1994). The (Non)Realization in Children's Utterances: Evidence for a Rhythmic Constraint. *Journal of Child Language, 21*, 59–83. DOI: <https://doi.org/10.1017/S0305000900008679>
- Wright, K., Poulin-Dubois, D., & Kelley, E. (2015). The animate-inanimate distinction in preschool children. *British Journal of Developmental Psychology, 33*, 73–91. DOI: <https://doi.org/10.1111/bjdp.12068>
- Yamamoto, M. (1999). *Animacy and reference. A cognitive approach to corpus linguistics. Studies in language companion series, 46*. Amsterdam, Philadelphia: John Benjamins Publishing Company. DOI: <https://doi.org/10.1075/slcs.46>

How to cite this article: Franz, I., Kentner, G., & Domahs, F. 2021 The impact of animacy and speech rhythm on the word order of conjuncts in German preschoolers and adults. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 12(1):12, pp.1–38. DOI: <https://doi.org/10.5334/labphon.254>

Submitted: 30 November 2020 **Accepted:** 05 April 2021 **Published:** 25 June 2021

Copyright: © 2021 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Laboratory Phonology: Journal of the Association for Laboratory Phonology* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS 