The Effect of Animacy and Noun Phrase Length on Austronesian Voice Production: An Experimental Approach

Eli Asikin-Garmager

1. Introduction and background

1.1. The effect of animacy and noun phrase (NP) length on word order

Speakers’ production patterns are clearly shaped by the structural properties of their specific language(s), but what about instances when multiple word orders and grammatical voice choices are available? When languages allow for syntactic options, what universal non-syntactic constraints influence word order? While the full inventory of biases universal to the production system is far from completely understood, one line of inquiry has been guided by Bock and Warren’s (1985:5) notion of Conceptual Accessibility (CA), or the “ease with which the mental representation of some potential referent can be activated in, or retrieved from, memory.” In respect to an example of heavy NP shift in English (1), CA would posit that the alternatives in (1) are not a matter of stylistic choice, but the preference to produce the heavier NP direct object utterance-finally (1b) – instead of before the ‘lighter’ (i.e., shorter) indirect object (1a) – reflects biases inherent to the human language production system.

(1) Heavy NP Shift in English
a. Tom gave the banana that the child chewed but subsequently vomited to her;
b. Tom gave her the banana that the child chewed but subsequently vomited.

In fact, the two objects in (1) differ not only in terms of heaviness, but also in terms of animacy, and previous investigations have examined the effects of animacy and NP length on syntactic production. Beginning with animacy, Jaeger and Norcliffe (2009) review several studies, including those of Bryne & Davidson (1985), Bock et al. (1992), Rosenbach (2005), and Bresnan et al. (2007). These studies show that speakers typically order animate NPs before those that are inanimate. As for NP length, previous work suggests that, at least for head-initial languages, there appears to be a universal short before long bias (Jaeger & Norcliffe, 2009), meaning that when two NPs in an utterance are not matched for length, speakers exhibit a bias to produce the shorter NP first. Yamashita and Chang (2001) provide a clear example of the importance of widening our typological lens, as their findings on length biases demonstrate the opposite preference among Japanese speakers. In contrast to English speakers, Japanese speakers exhibit a long before short bias, which Hawkins (2007) has argued is related to head-directionality.

As argued by Anand et al. (2011), Hawkins (2007), and Jaeger and Norcliffe (2009), there is a great need to incorporate more cross-linguistic data into the development of psycholinguistic theories, so in addition to improving our understanding of the Sasak language, the current investigation into the effect of animacy and NP length in Sasak can also help advance our broader understanding of universal production biases. As a preview, the Sasak data will provide additional evidence in favor of the animate before inanimate bias, but challenge claims that length biases are tied to a language’s head-directionality.

* Eli Asikin-Garmager, The University of Iowa, eli-asikin-garmager@uiowa.edu. I wish to thank William D. Davies and Emilie Deustrel Johnson for discussion of data and assistance with analysis. A sincere thanks to Nur Ahmad (University of Mataram), Titto, Risma Ayu, Atmimi Suciati, Mulyati, Alfian Juliansyah, Mustapa, Karin Selma Al-Kautsar, Ilyas, and many others for assistance with project logistics. My gratitude to all Sasak participants for their time and patience. Research was supported by a Fulbright-Hays Dissertation Research Abroad Fellowship from the US Department of Education, and a University of Iowa Graduate College T. Anne Cleary Research Fellowship.

properties. This is because, like English, Sasak is a head-initial language, but like Japanese speakers, Sasak speakers exhibit a long before short bias.

1.2. Sasak voice system and word order

Sasak is a Western Malayo-Polynesian language spoken primarily on Lombok Island (West Nusa Tenggara Province) in Eastern Indonesia by roughly 2.6 million people as of 2010 (Austin, 2012). Sasak’s grammatical voice system readily allows speakers to produce several different word orders, including both agent- and patient-initial orders (Asikin-Garmager, 2017). For example, clause-initial agents are canonically found in actor voice (AV) (2), while clause-initial patients are found in non-actor voice (NAV) (3). There are additional ways of arriving at agent- and patient-initial word orders. In addition to AV, agent-initial word orders are also found with antipassives (AP) (4), and patient-initial word orders are found not only with NAV, but also with passive voice (PV) (5). Again, Sasak’s grammatical voice system allows for multiple ways of arriving at both agent- and patient-initial word orders, with each voice type associated with a canonical word order. Due to space limitations, this paper will present the results of word order analyses in Section 2.

(2) kanak meme inó maléq bèmbéq inó  Actor voice (AV)
child male that AV.chase goat that
‘The boy chased the goat’

(3) bèmbéq inó paléq=ne, (isiq kanak meme inó)  Non-actor voice (NAV)
goat that NAV.chase=3 (by child male that)
‘He (the boy) chased the goat’

(4) terus acong=ne milu, ie me-mêta  Antipassive (AP)
then dog=3POSS follow 3 AP-search
‘Then his dog followed, continuing to search’

(5) (oku) wah=ku te-jeléq (siq lóq Ali)  Passive voice (PV)
(1SG) PERF=1SG PASS-hit (by ART.M Ali)
‘I was hit by Ali’

2. Production experiment
2.1. Participants and materials

Thirty-four participants (13 male and 21 female) were recruited for the current experiment. Half of the participants were from the Central (Menó-mené) dialect area and the other half were Ngenó-ngené (Easter dialect) speakers. Participants’ ages ranged from 18 and 35.

Design of materials and procedures is based partially on Yamashita and Chang’s (2001) Japanese production study examining the effect of NP length on word order. One important difference is that the current study includes animacy in addition to NP length as an independent variable. Five common transitive verbs were selected (chase, lift, blow, hit, and photograph). For each of these verbs, a set of NPs was designed to manipulate for animacy and length. Thus, each experimental item consisted of a transitive verb and two NPs (an agent and patient). There were three conditions for animacy and three for NP length. So, to construct each experimental item, first, agents and patients were selected for each verb. Each pair of NPs represented one of three pairings for animacy: (1) animate agent and patient, (2) animate agent and inanimate patient, and (3) inanimate agent and animate patient. Next, the agents and patients were modified for length. In the first condition the agent and patient NPs were equal in length. NP length was modified in the other two conditions by adding two adjectives to describe either the agent or patient NP. For convenience, these two additional conditions are termed long agent and long patient. Table 1 below illustrates the full range of experimental conditions (alongside word order predictions for each) for each lexical verb. Finally, while the agent and patient NPs were presented orthographically,

1 All examples presented in Section 1 come from Eastern (Ngenó-ngené) Sasak speakers. The following abbreviations are used in the glosses: AP – antipassive; ART – article; AV – actor voice; M – masculine; NAV – non-actor voice; PERF – perfective; POSS – possessive; PV – passive voice; SG – singular.
the verbs were presented via visual depictions (black and white images) because presenting them in either a prefixed or unprefixed form would have potentially implicitly biased participants production since each verb form (including the unprefixed form) is associated with a voice type (and thus word order).

2.2. Procedure and predictions

Participants performed a prompted recall picture description task. They sat in front of a laptop, which displayed a series of three screens for each experimental item (6). The first screen contained the visual depiction of a semantically transitive event corresponding to a pre-selected transitive verb (e.g., ‘chasing’ in (6)). Two NPs were presented in the upper corners of the screen, and these referred to the participants in the image. All visual and orthographic representations were counterbalanced. Participants viewed the image and NPs, but did not speak aloud during screen #1.

Table 1: Experimental conditions and word order predictions

<table>
<thead>
<tr>
<th>NP Length (3 conditions)</th>
<th>short agent short patient</th>
<th>long agent short patient</th>
<th>short agent long patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>short patient</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PVA</td>
<td>AVP</td>
<td>(conflict?)</td>
<td>AVP</td>
</tr>
<tr>
<td>long patient</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVP</td>
<td>Baseline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>short patient</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PVA</td>
<td>PVA</td>
<td>(conflict?)</td>
<td></td>
</tr>
<tr>
<td>long patient</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As we are investigating production biases, in no circumstance do we predict complete, 100% dominance of a single word order in any of these conditions. Instead, we anticipate analyzing the relative distributions of word orders when compared to a baseline condition. Table 1 presents the predictions for word order biases exhibited under different experimental conditions.

2 I present glosses and translations below for readability; only the Sasak text was presented to participants.

(i) *nine solah siq takut* | (ii) *mame*
girl beautiful REL scared | boy
‘The beautiful and scared girl’ | ‘The boy’

(6) Series of three computer screens for each experimental item

After viewing screen #1, participants pressed any key to advance to the next screen, which contained a simple math task (e.g., $15 - 8 = ?$). Participants answered the math task aloud, and then pressed any key to advance to the third (and final) screen for each item. Following Yamashita and Chang (2001:B48), the purpose of the math task was to provide a brief distractor and, “encourage subjects to produce their sentences from the meaning, rather than by covert rehearsal or reading from their visual buffer”. Finally, the third screen presented the same image as on screen #1, but this time without the NPs presented at the top. The participants were instructed to recall the words and phrases from the first screen and produce a single sentence describing the event depicted in the image. These utterances were recorded and transcribed.
(note: AVP = agent – verb – patient; PVA = patient – verb – agent). Because short NPs can more quickly be recalled from short-term memory (MacDonald, 2013) – or because Sasak is head-initial and speakers want to Minimize domains (Hawkins, 2007) – short agents and patients should occur in earlier linear positions. Thus, when the agent and patient differ in length, the shorter of the two should occur first. For example, if the shorter of the two NPs is the agent, we predict AVP. Next, given the proposed universal animate before inanimate bias, in the experimental condition with an animate agent and inanimate patient, for example, speakers should more frequently produce AVP utterances. As for the boxes with italicized word orders, the effects of animacy and length are expected to converge, resulting in even stronger biases for one word order. Predictions for the boxes marked ‘conflict’ are less clear.

2.3. Results

Following the data cleaning process, a total of 1187 utterances were analyzed. Coding for the dependent variable word order was collapsed into two categories: agent-initial and patient-initial. To examine the effect of the independent variables (i.e., animacy and length) on word order distribution, first a baseline comparison is needed (i.e., a default word order). To establish a baseline, I examined the distribution of agent- and patient-initial word orders when both animacy and length were matched. In this condition (presented in the middle of Figure 1), 62% of participants’ utterances were agent-initial and 38% were patient-initial.

![Figure 1: Distribution of agent- and patient-initial word orders across experimental conditions](image)

Examining the effect of animacy first, when the agent was animate (and patient inanimate), speakers’ production of agent-initial orders increased, and went from comprising 62% to 88% of total utterances. At the opposite end of Figure 1, the pairing of an animate patient with inanimate agent had the opposite effect. In this case, 88% of speakers’ productions exhibited patient-initial orders, while agent-initial orders decreased from comprising 62% (in the baseline condition) to only 12% of total utterances. As for the effect of length, when the agent was longer (relative to the patient; i.e., ‘agent [+long]’ condition), there was an increase in the production of agent-initial orders (an increase from 62% to 69%). Conversely, in the long patient condition, the production of agent-initial word orders decreased markedly, and patient-initial orders went from comprising 38% to 54% of total utterances.

Given the binary nature of the dependent variable, to statistically assess these effects, a binomial logistic mixed-effects model was implemented with the lme4 package (Bates et al., 2015) and glmer function in the statistical environment R (R Development Core Team, 2016). Unless otherwise noted, all models reported here included a random intercept for participant and item number. Both factors animacy and NP length were included as fixed effects, and I adopt an alpha level of \( p < .05 \) for statistical significance. The presence of an animate agent resulted in a highly significant increase in the odds of observing an agent-initial word order (\( \beta = -1.3854, SE = 0.2624, p < .001 \)). In contrast, the presence of an animate patient had the opposite effect, resulting in significantly greater odds of speakers producing patient-initial utterances (\( \beta = 2.5203, SE = 0.2810, p < .001 \)). As for NP length, a long patient had the effect...
of significantly increasing the odds of speakers producing a patient-initial utterance ($\beta = 0.9136$, $SE = 0.2760$, $p < .001$). Finally, while a long agent resulted in a marked increase in the number of agent-initial utterances, this effect was statistically insignificant ($\beta = -0.4049$, $SE = 0.2680$, $p = 0.130$).

Overall, we find that the effect of animacy was greater than that of length, although there was also a statistically significant effect of length in the long patient condition. To more closely examine the effect of length, I examined two subsets of the data.

First, I examined the subset in which the agent was animate and patient inanimate. As we already know, animate agents strongly bias agent-initial word orders, but what happens as NP length is also manipulated? As shown in Figure 2, the lengthening of an agent had little to no effect on the likelihood of observing an agent vs. patient-initial order. Conversely, when the patient was longer than the agent, there was a marked decrease in the number of agent-initial word orders and increase in number of patient-initial orders. The number of utterances with agent-initial word order fell from comprising 88% to only 72% of all utterances, while the number of patient-initial orders increased from comprising 12% to 28% of all utterances. This suggests that while an animate agent biases an agent-initial word order, a long patient mitigates this effect.

![Figure 2: Effect of length on word order when agent is animate](image)

To statistically assess the changes presented in Figure 2, I conducted a model on the subset of data in which the agent was animate and patient inanimate. There was no significant effect of a long agent ($\beta = 0.1924$, $SE = 0.7437$, $p = 0.795$). However, there was a statistically significant effect of a long patient ($\beta = 1.6465$, $SE = 0.7364$, $p = 0.025$), meaning that a long patient resulted in greater odds of observing a patient-initial word order.

Next, I examined the effect of NP length on the subset of data in which the patient was animate and agent inanimate (Figure 3). As already noted, animate patients strongly bias a patient-initial word order.

![Figure 3: Effect of length on word order when patient is animate](image)

As shown in Figure 3, in the long agent condition, there was a decrease in the number of patient-initial orders. Patient-initial orders decreased from comprising 88% to 76% of all utterances. Conversely, agent-initial orders went from comprising 12% to 24% of all utterances, doubling in frequency. This
suggests that the long agent may have mitigated the effects of the patient-initial bias in the animate patient condition. Next, when the patient was not only animate, but also longer than the agent, we observed a marked increase in the use of patient-initial utterances. Patient-initial orders increased from comprising 88% to 95% of all utterances. So, while an animate patient already biases a patient-initial word order, when the patient is also longer, the number of patient-initial orders is even greater, suggesting a facilitatory effect of length on animacy.

I conducted a model to statistically assess the effect of length on the subset of data in which the agent was inanimate and patient animate. In this case, there was a main effect of length. In the presence of a long agent, the odds of observing an agent-initial utterance increased significantly ($\beta=-0.9882, SE=0.3800, p=0.009$), and the presence of a long patient had the opposite effect ($\beta=1.2747, SE=0.5608, p=0.023$), resulting in greater odds of speakers producing a patient-initial utterance.

In summary, results show a strong animate first bias. If the agent is animate, we are more likely to observe an agent-initial order, and animate patients bias patient-initial word order. As for length, when the agent and patient are matched for animacy (i.e., both animate), the presence of a long patient NP strongly biases a patient-initial ordering, demonstrating a long before short bias. Note that this is opposite of what was predicted. The effect of a long agent was also significant, but only when its effect was not washed out by the already strong effect of animacy in biasing agent-initial word orders. We see this by comparing the effect of the long agent in Figures 2 and 3.

3. Discussion and the Primacy of Semantic Richness before Event Structure

While the animate before inanimate bias observed among Sasak speakers provides further evidence in support of its universality, given that Sasak speakers also exhibited a long before short bias, we must reject the hypothesis that shorter NPs are universally given sequential priority in the architecture of the production system. Moreover, because Sasak is a head-initial language, yet patterns with a head-final language (i.e., Japanese) in terms of NP length effects, we cannot attribute a long before short or short before long bias to head-directionality. As such, the resulting challenge is to account for different length biases in different languages.

Important to the account proposed here is Yamashita and Chang’s (2001) observation that longer NPs have conflicting characteristics. Because long NPs are long, they are more difficult to recall and produce, given their greater load on working memory. On the other hand, the length of long NPs also means they are semantically rich and more cognitively salient (e.g., James et al., 1973; as cited in Bock & Warren, 1985). Next, note how in addition to retrieving words during sentence production, speakers must also formulate structural relationships and dependencies between words and linearize them over a very short course of time. Crucially, these structural dependencies increase significantly at the production of the lexical verb, which introduces event structure, in some cases tense and aspectual information, and structural biases associated with that verb (e.g., Colleman, 2009; Stallings et al., 1998; van de Velde & Meyer, 2014). For example, Stallings et al. (1998:396) demonstrate how verbs are accompanied by, “frequency-weighted lexical information.”

The hypothesis presented in (7) draws on the conflicting characteristics of long NPs while taking into account how production challenges change markedly over the course of sentence production with the demands accompanying the production of lexical verbs. The hypothesis in (7) not only is able to account for the current cross-linguistic NP length puzzle, but also makes clear empirical predictions.

---

3 The hypothesis is also consistent with Yamashita and Chang’s (2001) final note that different patterns may be observed pre- and post-verbally. Much like the account here, these authors (p.B54) argue similarly that, “form-related factors should differentially influence these [pre- and post-verbal] regions.” One important difference between these authors’ discussion and the account proposed here is how they differ in terms of universality. While Yamashita and Chang also suggest that the position of the verb relative to the observed shift is important, they still attribute the length effects to language-specific differences, when they state, “If the shifts in human languages are motivated by both conceptual and lexical factors, then the language-specific differences in sensitivity to these two factors can provide a uniform account [emphasis added] for both “short before long” and “long before short” preferences.” In contrast, I maintain that the different length effects are universally observed as a function of where the shift occurs in relation to the verb. While in a strictly verb-final language, for example, this distinction becomes opaque since the shift can only occur in a preverbal position, any language in which the shift could occur either pre-or post-verbally clarifies why this account is not tied to language-specific differences.
(7) **Primacy of Semantic Richness before Event Structure Hypothesis**

The semantic richness and cognitive salience of a long NP takes precedence over its increased demand on working memory at any temporal/linear position prior to the lexical verb.

The hypothesis in (7) captures how before the production of the lexical verb, the semantic richness of long NPs takes precedence over their increased demands on cognition and production, and speakers thus exhibit a bias to produce long NPs before relatively shorter ones. However, as production of a sentence proceeds incrementally, the increasing demands of formulation and execution reach a threshold with the introduction of the verb. At this point, speakers begin to exhibit a bias to produce shorter NPs before relatively longer ones due to the decreased demands that shorter NPs exert on working memory and their relative ease of production. As such, we observe short before long biases preverbally, and long before short biases post-verbally.

The advantages of the account proposed here are that it (1) most importantly, accounts for the otherwise varied length biases exhibited by speakers of different languages, (2) avoids evoking language-specific constraints or generalizations where otherwise unnecessary, and (3) makes several empirical predictions. For example, it predicts that languages cannot be categorized into binary long before short or short before long categories since whether speakers exhibit one bias or the other is a function of where the shifted domain is relative to the lexical verb. For example, if speakers of a so-called short before long language (e.g., English) were given a task in which NPs competed for a sentence-initial position, the account predicts that the longer NPs would be given linear precedence due to their semantic richness. Yamashita and Chang (2001) discuss one case of both length biases being observed in a single language, noting Hawkins’ (1994:131) discussion of how Hungarian speakers exhibit a short before long bias post-verbally, but order longer NPs first in the preverbal domain.

The hypothesis in (7) also makes clear predictions for verb-initial languages. As the domain for any NP shifts is exclusively post-verbal in these languages, the **Primacy of Semantic Richness before Event Structure Hypothesis** predicts that all length effects should fall into the short before long category. Māori exhibits both VSO and VOS word order, and Chung (1998:170-171; as cited in Davis (2005:58)) discusses how there is a preference for pronouns before nonpronouns. One distinguishing characteristic of pronouns is, of course, that they are shorter and contain less phonological material. In a separate source, the role of a short before long preference in Māori post-verbal linearization is noted by Bauer (1993:245; as cited in Davis (2005:58)). Next, in Tagalog actor voice, both VSO and VOS orders are possible, and Kroeger (1993b:111; as cited in Billings (2005:108)) discusses a weight shift – namely that heavy-NPs are shifted to later sentence positions. While further work isolating the effect of length on word order in verb-initial languages is needed, there is clear initial evidence that predictions of the current account are borne out.

4. **Summary**

This paper reported results from a Sasak language production experiment, which demonstrate that speakers’ word order choices are shaped by both the animacy and relative length of verbal arguments in the clause. Specifically, Sasak speakers exhibit an animate before inanimate, as well as a long before short, bias. As such, the resulting puzzle is that data from some languages (e.g., English, German) reveal a short before long bias, while data from other languages (e.g., Sasak, Japanese) demonstrate a long before short bias. This paper argues that these different biases are not language-specific, but fallout from the **Primacy of Semantic Richness before Event Structure Hypothesis**, which states that ‘the semantic richness of a nominal element takes precedence over its increased demand on working memory at any temporal/linear position prior to the lexical verb.’ In other words, length effects can be understood as a function of where the shifted NP domain is relative to the lexical verb because of the demands that the verb places on form-oriented production processes. Foremost, the hypothesis accounts for a range of cross-linguistic data, and does so without stipulating language-specific parameters. Moreover, it makes several clear empirical predictions, one of which is supported by data from verb-initial languages.
References


