Speech Rates Differentiate Nouns and Verbs in Child-Surrounding and Child-Produced Speech: Evidence from Chintang

Nicholas A. Lester, Balthasar Bickel, Steven Moran, and Sabine Stoll

1. Introduction

One of the earliest challenges any child experiences when learning their first language is to differentiate the various syntactic functions of words. Several factors are known to play a role in this process, including referential semantics (Pinker, 1984), syntactic distributions (Landau & Gleitman, 1985; Naigles & Swenson, 2007; Fisher, Gertner, Scott & Yuan), segmental distributions (e.g., Saffran, Aslin, & Newport, 1996) and lexical environments (e.g., Redington, Chater, & Finch, 1998; Mintz, 2003; Moran, Blasi, Schikowski, Küntay, Pfeiler, & Stoll, 2018). The present paper explores how an additional component, prosody, plays a role. Specifically, we focus on whether and how speech rates distinguish nouns from verbs in child-surrounding and child-produced speech in a language hitherto unexplored: Chintang.

Prosody has been shown to distinguish nouns and verbs in adult-directed speech. Key evidence comes from studies of homophones with a noun or verb reading, as in English *fish*-NOUN vs. *fish*-VERB. Sorenson, Cooper, & Paccia (1978) had participants produce sentences that contained either the noun or verb variant of English homophones. They manipulated position within syntactic units (phrase-initial vs. phrase-final), while controlling for phonetic environments, stress patterns, and other factors. When variants were allowed to occupy different phrasal positions, nouns (phrase-final position) were produced with longer durations than verbs (phrase-initial position). However, this effect disappeared when both nouns and verbs were placed in the same syntactic-phrasal positions.

---

[*] Nicholas A. Lester, Department of Comparative Language Science (CLS) and Center for the Interdisciplinary Study of Language Evolution (ISLE), University of Zurich, nicholasandrew.lester@uzh.ch. Balthasar Bickel, CLS and ISLE, University of Zurich. Steven Moran, CLS and ISLE, University of Zurich. Sabine Stoll, CLS and ISLE, University of Zurich. This research is part of the project ‘Acquisition processes in maximally diverse languages: Min(ding) the ambient languages (ACQDIV)’ that has received funding from the European Research Council (ERC) under the European Union’s Seventh Framework Programme (FP7-2007-2013) (Grant agreement No. 615988; PI Sabine Stoll). NAL, BB, SS: developed concept and research design. NAL: performed the research and statistical analyses. BB, SM, SS: provided data. NAL, BB, SM, SS wrote the paper.

Proceedings of the 44th Boston University Conference on Language Development, ed. Megan M. Brown and Alexandra Kohut, 280-293. Somerville, MA: Cascadilla Press.

They conclude that syntactic-phrasal position is the primary determinant of length differences between word classes in English.

Lohmann and Conwell (2020) report additional factors that interact with word class and duration. They take advantage of the fact that the strength of prosodic boundaries correlates positively with duration (e.g., Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992), while token frequency correlates negatively (Gahl, 2008). In their sample, nouns were both more likely to occur before a stronger prosodic boundary and to have lower token frequency. These two factors accounted for independent portions of the variance in duration, and when considered jointly, virtually eliminated any trace of a category-specific effect.

One feature of these studies is that in pursuit of careful phonological control, they only consider homophonic pairs. This approach works for languages with shallow morphology like English (Conwell, 2017; Conwell & Morgan, 2012; Lohmann & Conwell, 2020) and Chinese (Li, Shi, & Hua, 2010), or languages with ambiguous morphophonology, like French (Shi & Moisan, 2008). However, it does not apply to languages in which words bear a greater degree of category-specific inflectional morphology, such as Chintang.

Others have examined speech rates immediately prior to nouns and verbs. Research on lexical production suggests that planning from concept to articulation takes approximately 600 ms (Indefrey & Levelt, 2004). Seifart et al. (2018) sample speech rates within this window before nouns and verbs. They find that speakers of many typologically distinct languages slow their speech more before nouns than before verbs, mirroring the work on noun durations. These speakers were also more likely to pause or become otherwise disfluent before nouns than verbs. Both slower speech rates (e.g., Gahl, 2008) and disfluency (Rochester, 1973) are associated with increased cognitive effort, suggesting that the noun/verb distinction is in part related to psycholinguistic factors beyond the prosodic structure of the utterance. Accordingly, Seifart and colleagues speculate that these differences may arise due to difficulties associated with higher-order processes of noun use, namely pragmatic principles of reference and the incorporation of new information.

Prosody has previously been demonstrated to play a strong role in language acquisition. Infants as young as 9 months use prosodic cues to segment words (Johnson & Jusczyk, 2001; Mattys, Jusczyk, Luce, & Morgan, 1999) and larger syntactic units (Christophe, Millotte, Bernal, & Lidz, 2008; Hirsh-Pasek, Kemler Nelson, Jusczyk, Cassidy, Druss, & Kennedy, 1987; Kemler Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989), to track and incorporate grammatical morphemes during sentence comprehension (Gerken & McIntosh, 1993), and to discriminate between languages (Nazzi, Bertonocini, & Mehler, 1998). Very young children are therefore quite capable of using prosodic features of the input when learning various aspects of the language that surrounds them.

Other work shows that nouns and verbs are prosodically distinguished in child-directed speech as with adult-directed speech. For example, Conwell (2017) extends the experimental results of Sorenson et al. (1978) to observational data on English child-directed speech. She finds that words that can be ambiguous between noun and verb categories are produced with longer durations when they occur as nouns (see also Conwell & Morgan, 2012).

Less is known about these issues in child speech, or whether children indeed use prosodic information about word class as they learn their language. Some evidence for the development of prosodic features comes from school-age children (kindergarten through second grade). Barth (2015) finds that children and adults modulate speech rates according to the same variables. Specifically, they produce function words with shorter durations in low-information contexts. The same study showed that children gradually approximate the durations observed for adults, and that this effect could not be attributed solely to increased motor-articulatory skills. Therefore, at least some aspects of children's prosodic development appear to be learned in a domain-specific fashion. However, we still do not know whether children distinguish content words by prosodic means in their own speech, nor how such prosodic asymmetries develop over time.

The present study draws these different threads of research together to examine the relationship between prosody and syntactic class in the speech of young children and adults. We ask the following three questions. First, does the finding that nouns have longer durations than verbs in English child-directed speech extend to inflectionally much richer languages? To this end, we test whether nouns and verbs in Chintang, a morphologically rich language, are distinguished by duration when other relevant factors have been controlled for. Our second question asks whether differences in speech rates immediately before nouns and verbs extend from adult-directed speech to child-directed speech. Differences in durations at the target word have been shown to surface in both adult-directed and child-directed speech, despite well-documented prosodic differences between the two registers overall (Conwell, 2017; Li, Shi, & Hua, 2017; Lohmann & Conwell, 2020; Shi & Moisan, 2008; Sorensen et al., 1978). However, no study has examined this possibility in preword windows, which have been hypothesized to depend more heavily on higher-order processes in speech production, such as information management. The two kinds of effects may play out differently in speech to children than in speech to adults. Finally, we ask whether the relative prosodic differences between nouns and verbs differ in the speech of Chintang children and adults.

2. Data and Methods

2.1. The Chintang language

The data for this study are taken from a longitudinal corpus of Chintang language acquisition (Stoll et al., 2015), included as part of the ACQDIV database (Moran, Schikowski, Pajović, Hysa, & Stoll, 2016). Chintang is a polysynthetic Tibeto-Burman language spoken in a village in Eastern Nepal. It is currently considered endangered according to UNESCO standards for evaluating linguistic vitality (UNESCO Ad Hoc Expert Group on Endangered Languages, 2003). The basic sentence structure is SOV. Lexical nouns and verbs are typically multimorphemic. Nouns inflect for case, number, and possession (Schikowski, Paudyal, & Bickel, 2015). Verbs inflect for tense, aspect, mood, and polarity (order is not entirely fixed), and exhibit double agreement; further, they may combine with secondary stems (Bickel et al., 2007). Thus, Chintang verbs are morphosyntactically much more complex than nouns.

2.2. Corpus

The Chintang corpus is made up of audio-visual recordings of six target children ranging in age from 0;5.2 at the youngest to 4;4.29 at the oldest. An additional 155 adults also participate in the recordings. These adults range in age from 13 to 80 years. Recordings were made in naturalistic settings with no intervention from the researchers. Interactions among many different interlocutors were captured for each child, including family members, but also unrelated children and adults from the village. These data were then transcribed and morphologically annotated by trained experts. The entire corpus contains 987,672 words of child-produced and child-surrounding speech. During transcription, recordings were segmented into utterances, which we refer to as *annotation units*. These are the foundation for further acoustic analysis. Descriptive statistics for the overall corpus are presented in Table 1.

Table 1: Descriptive statistics of the Chintang sample

	Child speech	**Child-surrounding speech**
Noun types	19,083	4,945
Nous tokens	134,009	26,885
Verb types	35,107	7,389
Verb tokens	166,739	23,000

2.3. Data preparation

First, we convert the orthographic transcriptions to phonetic transcriptions using the mapping techniques presented in Moran and Cysouw (2018). This step is necessary, as the forced-alignment procedure described below relies on matching phonetic labels to acoustic signals. Once we have the phonetic transcription and the audio, we can align the two automatically via WebMAUS (Kisler, Reichel, & Schiel, 2017). We also extract the rates of speech immediately prior to the target word using the same technique. Following Seifart et al. (2018), we aim for a preword window of 500 ms (see above).

2.4. Measures

We compute two measures of the speech rate: one for the target word and one for the preword window. Speech rate was calculated as the number of segments divided by the length of the word or preword window in seconds. We further collect information about the following variables: speaker, session, target word form, initial segment, phonological length, position of the word within the annotation unit, word class of word in preword window[1] (*noun, verb, other*), word class of target (*noun, verb*), speaker age (*adult, child*).

Speaker, session, and target word form are collected to account for idiosyncratic features of each given that our data consist of multiply interdependent samples. Initial segments are more or less easily synced from transcription to audio, depending on their phonetic properties.. For example, the voice-onset time of initial voiceless obstruents creates problems for estimates of word onsets, while sonorant onsets are more reliably locked to the signal given their more robust acoustic signature. Phonological length is included because it is known to correlate with speech rates (Lehiste, 1970) and potentially also with production onset latencies in isolated word production (Alario et al., 2004). Likewise with position of the word in the utterance, which is associated with duration (e.g., phrase-final lengthening; Turk & Shattuck-Hufnagel, 2007), and modulation of duration based on syntactic category (e.g., Sorensen et al., 1978). For the speaker age variable, adults were defined as speakers aged 13 years or more at the time of recording. Child refers only to the target children (i.e., we exclude non-target children under the age of 13). We do this for two reasons: to ensure a greater number of observations per child (many non-target children are sampled only sparsely), and to concentrate our child sample on the youngest age group (rather than ranging from six months to 12 years). Target rates (when applied as a predictor variable) and word lengths were normalized using z-scores, defined as $(x_i - \overline{x})/\sigma_{\overline{x}}$, where x_i is the length of a given token, $\overline{x}$ is the mean length for the entire sample, and $\sigma_{\overline{x}}$ is the standard deviation of the sample. Positions within the annotation unit were normalized to a percentage scale [0,1] to account for their variable lengths. Normalized position is defined as $(i-1)/(n-1)$, where i is the offset from the beginning of annotation unit (starting at 1) and n is the length of the unit in words.

3. Results

We fitted two linear mixed-effect models. The first model predicts the speech rate of the target word for nouns and verbs. It therefore answers our first question: whether prosodic differences are observed between nouns and verbs for child-surrounding speech in morphologically complex languages. The second model predicts speech rates in the preword windows. It therefore answers our second question: do preword asymmetries in speech rate between nouns and verbs appear in child-directed speech? Both models include a predictor distinguishing speaker age, allowing us to address our third question concerning whether and how child-surrounding speech relates to child-produced speech.

1. Our sample contained only a single word within all preword windows.

3.1. Speech rate at target

To ensure comparability between this analysis and the one predicting speech rates in preword windows, we only include target words that are preceded by some content (i.e., not only a pause) within the annotation unit. This trim leaves us with 132,208 tokens across nouns and verbs. An initial visual inspection of the speech rates revealed a strong positive skew. In order for the assumptions of the linear model to be met, the dependent variable must approximate a normal distribution. A Box-Cox power analysis (Box & Cox, 1964) suggested that the target rates would better approach normality if we apply a square-root transformation ($\lambda \sim 0.5$). Visual inspection confirmed the success of this transformation.

Fixed effects include the *z*-scored length of the target word, the normalized position of the word within the annotation unit, speaker status (*adult* or *child*), the initial segment of the target word, and its part of speech (*noun* or *verb*). We further include the interaction of speaker status with all variables except for initial segment, which was excluded due to issues of data sparsity. Finally, we include random intercept adjustments for individual speakers, sessions, and target words. This model is isomorphic to the ones reported in Seifart et al. (2018), with the exception of the additional age-related interaction terms, as well as the main effect of initial segment.

Table 2 provides a summary of the model. Columns reflect the model coefficients β (in the square-root-transformed scale), standard errors, degrees of freedom, *t*-values, and *p*-values. Coefficients are modeled using sum-to-zero contrasts. Levels of categorical variables are ordered alphabetically. Coefficients for initial segment are omitted for reasons of space, though this predictor was significant ($F(28, 10324) = 20.43$, $p < .001$).

Table 2: Summary of model predicting speech rate at the target.

Random effects	**Variable**	**Variance**	**SD**		
	word	0.10	0.32		
	session	0.08	0.28		
	speaker	0.02	0.13		
	Residual	**0.86**	**0.93**		
Fixed effects	***β***	**SE**	**DF**	***t***	***p***
intercept (grand mean)	3.92	0.04	221.90	98.72	<. 001
speaker age (adult)	0.23	0.03	87.62	7.45	< .001
word class (noun)	0.01	0.01	38020.00	1.72	.09
position	-0.44	0.01	131800.00	-29.93	< .001
phon. length	0.02	0.01	47750.00	4.60	< .001

speaker age (adult) × word class (noun)	0.00	0.00	127300.00	0.92	.36
speaker age (adult) × position	0.05	0.01	131500.00	3.63	< .001
speaker age (adult) × phon. length	0.01	0.00	117300.00	1.43	0.15

In addition to initial segment, the other control variables were significant and in the expected direction. Regarding utterance position, nouns and verbs are produced with slower speech rates at later positions ($F(1, 131768) = 895.81$, $p <$ .001.). This effect is modulated by speaker age, such that children slow down to a greater degree than adults ($F(1, 131523) = 13.20$, $p < .001$.). Phonological length was also a significant predictor of target rates ($F(1, 47755) = 21.16$, $p < .001$.): longer words are produced with slightly faster rates (a phoneme-by-length compression effect; e.g., Lehiste, 1970). This effect does not interact with speaker age.

Our critical variables were speaker age and word class, along with their interaction. Our model revealed a highly significant main effect of speaker age ($F(1, 88) = 55.56$, $p < .001$). The positive coefficient indicates that adults are generally faster than children. However we fail to find either a reliable main effect of word class ($F(1, 38025) = 2.96$, $p = .09$) or of the interaction between word class and speaker age ($F(1, 127279) = 0.85$, $p = .36$). The relationship between word class and age group is plotted in Figure 1. Note that, if anything, nouns are produced slightly *faster* than verbs by both age groups (also indicated by the positive overall coefficient for nouns in Table 2).

3.2. Speech rate in the pre-word window

We next analyze speech rate in the pre-word window. The data for this analysis are identical to those from the previous model. All of the fixed and random effects from that model are included. We further include fixed effects of target speech rate, word class in preword window, and the interactions of both of these variables with speaker age. A Box-Cox power analysis likewise revealed the need for a square-root transform to the preword speech rates prior to modeling.

The resulting model is summarized in Table 3. Again, we exclude results for the initial segment for purposes of space, though the effect was also significant ($F(28, 10324) = 20.43$, $p < .001$). First, we assess the control variables related to the target word. Position was negatively correlated with speech rates in the preword window ($F(1, 131534) = 337.13$, $p < .001$). Nouns and verbs that occur later in utterances tend to be preceded by slower speech. This effect differs between adults and children similar to what we observed for target durations: children respond more strongly to position-based deceleration.

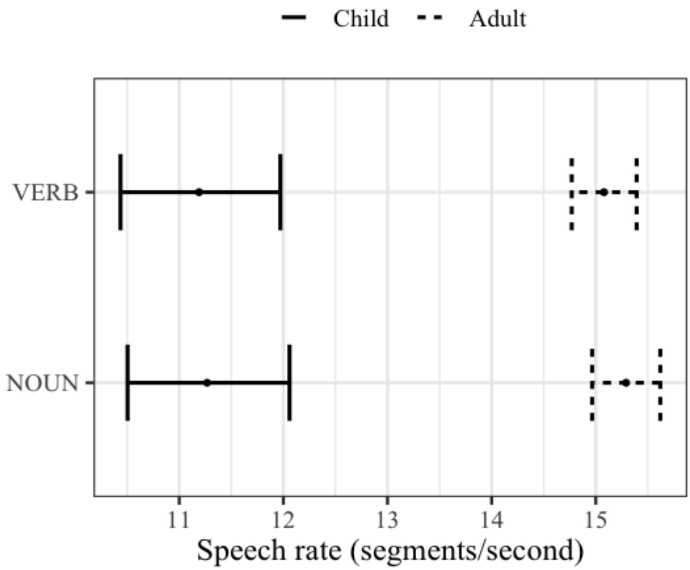

Figure 1: Fitted estimates of mean speech rates for nouns and verbs in child speech (Child) and child-surrounding speech (Adult). Whiskers represent confidence intervals.

Table 3: Summary of model predicting speech rate in the pre-word window

Random effects	**Variable**	**Variance**	**SD**		
	word	0.03	0.18		
	session	0.05	0.22		
	speaker	0.01	0.09		
	residual	**0.92**	**0.96**		
Fixed effects	***β***	**SE**	**DF**	***t***	***p***
intercept (grand mean)	4.16	0.05	1341	89.58	< .001
speaker age (adult)	0.05	0.02	101	2.17	< .05
word class (noun)	-0.03	0.01	18710	-6.73	< .001
position	-0.27	0.01	131500	-18.36	< .001
phon. length	0.07	0.00	29980	14.37	< .001
preword type (noun)	0.00	0.01	130400	0.02	0.99
preword type (verb)	0.06	0.01	130700	8.24	< .001

target rate	0.05	0.00	130600	12.98	< .001
speaker age (adult) × word class (noun)	0.00	0.00	107500	-0.81	0.42
speaker age (adult) × position	0.16	0.01	131800	10.75	< .001
speaker age (adult) × phon. length	-0.02	0.00	109500	-4.01	< .001
speaker age (adult) × preword type (noun)	-0.02	0.01	131500	-3.92	< .001
speaker age (adult) × preword type (verb)	0.01	0.01	131600	1.46	.14
speaker age (adult) × target rate	0.08	0.00	132000	20.17	< .001

Phonological length correlated positively with speech rates ($F(1, 29981) = 206.40$, $p < .001$). Longer nouns and verbs tended to be preceded by faster speech rates. Lengths of target and pre-target forms are significantly and positively correlated ($\rho = .10$, $p < .001$). Thus, this effect may stem from the same compression discussed for the previous model. However, it is important to note that the target effect is weaker and more variable than the preword effect (as evidenced by coefficients and standard errors). Thus, there are likely to be other factors at play, such as end-weight phenomena (e.g., Wasow, 1997). Longer units tend to occur after shorter units; they should therefore have naturally slower speech rates given that they occur later in the utterance. The effect of length interacted with speaker age, with adults showing a weaker trend ($F(1, 109507) = 16.10$, $p <.001$).

Target rate correlated positively with preword rate ($F(1, 130562) = 168.54$, $p <.001$), most likely reflecting an overall speech rate effect (faster overall speech rates manifest in targets and the words that immediately precede them). Again we find an age-related interaction: adults show a stronger relationship than children ($F(1, 131974) = 406.69$, $p < .001$).

The word class of the preceding word also proved significant ($F(2,130923) = 70.33$, $p <.001$). While preword nouns agree generally with the overall mean, verbs produce a slight increase in preword rates. However, adults and children differ in this regard. Adults slow preword rates more than children when the window contains nouns than verbs ($F(2,131666) = 9.53$, $p <.001$). However, this result should be interpreted with some caution, as we have not controlled for other aspects of the words in the preword windows (e.g., length, utterance position, and so on) as we did in the analysis of target rates, which did not reveal such a difference for Chintang in either adult or child speech.

Our critical variables are again speaker age, word class of the target, and the interaction between the two. The model uncovered a main effect of age: adults spoke at faster rates than children ($F(1, 101) = 4.69$, $p < .05$). Target word class was also significant as a main effect ($F(1, 18706) = 45.28, p < .001$): nouns slowed speech more so than verbs in the preword window. Crucially, this difference did not depend on age group ($F(1, 107463) = 0.66, p = .42$). Fitted estimates of preword rates for nouns and verbs in the speech of children and adults are plotted in Figure 2.

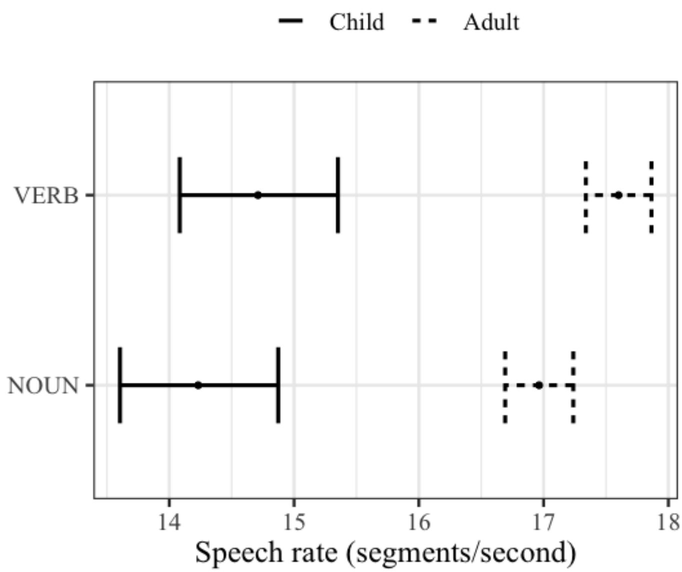

Figure 2: Fitted estimates of mean speech rate in the preword window for nouns and verbs in child speech (Child) and child-surrounding speech (Adult)

4. Discussion

We explored three questions about how speech rate interacts with word class in the productions of children and adults. First, we asked whether nouns and verbs are produced with different speech rates in child-surrounding speech in Chintang, a polysynthetic, Tibeto-Burman language that differs widely from the other languages that have been studied in similar paradigms (e.g., English, French, and Mandarin). Second, we asked whether speech rates differ between nouns and verbs in the window of time immediately before they are produced in child-surrounding and child-produced speech. This question extends work on adult-directed speech to a new register. Third, we asked whether the answers to the prior two questions hold in the speech of children as well as adults speaking in the presence of children.

Regarding the first question, we did not find any difference in the speech rates of nouns and verbs when other relevant factors were controlled for. We therefore do not replicate the findings for English, French, and Mandarin that have been reported in prior literature. This finding is all the more striking given that Chintang verbs are much longer and more complex than nouns. Increased length typically results in phonetic compression (i.e., faster rates for words with more segments; Lehiste, 1970). While we find that the mean rates for verbs were numerically faster than for nouns, the difference was not reliable. While higher resolution research is needed to fully assess this observation, the surprisingly slow speed of verb articulation can potentially be explained by the sheer complexity of Chintang verb morphology (Bickel et al. 2007).

Turning to the second question, we replicated the preword effect from the study of Chintang adult-directed speech included in Seifart et al. (2018): nouns in adults' child-surrounding speech likewise slow speech more so than verbs in the preword window. Therefore, the preword window effect appears to be independent of this contrast in register. This interpretation comes with a caveat, however: our corpus contains not only speech that is directed to children, but also speech that is directed to other adults who are coincidentally in the presence of children. Therefore, the lack of difference that we observe could be influenced by the presence of adult-directed speech. Nevertheless, it is worth noting that the magnitude of the word class effect observed here, as well as the estimated means for nouns and verbs, comports well with those reported for Chintang adult-to-adult speech (Seifart et al., 2018).

Finally, we find that the children in our sample generally behave like the adults with respect to our critical variables. Adults did not show any difference at the target between nouns and verbs, and neither did the children. Conversely, adults did show a difference in speech rates at the preword window, and so did the children. This suggests that the difference is not gradually learned from adult behavior but is present from the outset. If so, then where does it come from? In agreement with Seifart et al. (2018), we propose that prosodic characteristics of the preword windows for nouns and verbs arise naturally from fundamental features of the production architecture, namely information management. Introducing or contrasting nominal content, particularly in pro-drop languages like Chintang, is simply more taxing, especially during planning. Otherwise, the form could be omitted entirely, in which case the speaker could rely on agreement morphology in the verb. Thus, the effort of producing the form reflects its status as new, less-accessible, or generally more difficult to integrate with the existing discourse.

Together, these findings suggest that speech rates behave differently at the target than in the preword window. That is, while word class is strongly signaled by the preword window, there is no discernible difference at the target. This suggests that the two may stem from different processes. On the other hand, we found that word class does affect speech rates in the preword window, irrespective of the class of target word. Preword nouns slowed speech more than preword verbs. Because nouns and verbs do not appear to be differentiated by speech rates themselves, this difference must have another cause. One thing to consider is that preword speech rates include pauses (when present) anywhere within that window, at least as long as a word is present. By contrast, rate measurements at the target words were based solely on the segmental content from the onset to the end of a word. If pauses do indeed play a role, then this observation would be in line with the findings of Seifart et al. (2018), who found that nouns correlated with increases in the likelihood of pauses (something that we have not analyzed here).

We see at least two additional avenues for further research. First, the analyses reported here treat age as a dichotomous variable (target children vs. adults). The next step is to model age as a continuous variable, allowing us to get a more nuanced perspective on the development of prosodic noun/verb contrasts over time. Second, our general approach should be extended to more typologically diverse languages. We have already observed different effects in Chintang compared to those which have been reported for other languages (English, Mandarin, and French. The most important next step is to determine the scope of variability, as well as how this variability might be linked to typological features. The results of such an analysis will give us a better understanding of the mechanisms behind these contrasts and their ontogeny.

References

Alario, F.-Xavier, Ferrand, Ludovic, Laganaro, Marina, New, Boris, Frauenfelder, Uli H., & Segui, Juan. (2004). Predictors of picture naming speed. *Behavior Research Methods: Instruments & Computers*, *36*(1), 140-155.

Barth, Danielle Gilberte. (2015). To *have* and to *be*: Function word reduction in child speech, child directed speech and inter-adult speech (Unpublished doctoral dissertation). University of Oregon, Eugene, Oregon.

Bickel, Balthasar, Banjade, Goma, Gaenszle, Martin, Lieven, Elena, Paudyal, Netra Prasad, Ichchha, Purna Rai, Rai, Manoj, Rai, Novel Kishore, & Stoll, Sabine. (2007). Free prefix ordering in Chintang. *Language*, *83*(1), 43-73.

Boersma, Paul. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*, 341-345.

Box, G. E. P., & Cox, D. R. (1964) An analysis of transformations. *Journal of the Royal Statistical Society*, *Series B (Methodological)*, *26*, 211-252.

Christophe, Anne, Millotte, Séverine, Bernal, Savite, & Lidz, Jeffrey. (2008). Bootstrapping and lexical acquisition. *Language and Speech*, *51*, 61-75.

Conwell, Erin. (2017). Prosodic disambiguation of noun/verb homophones in child-directed speech. *Journal of Child Language*, *44*(3), 734-751.

Conwell, Erin, & Morgan, James, L. (2012). Is it a noun or is it a verb? Resolving the ambicategoriality problem. *Language Learning and Development*, *8*, 87-112.

Finch, Steven, & Chater, Nick. (1992). Bootstrapping syntactic categories. In *Proceedings of the Fourteenth Annual Meeting of the Cognitive Science Society* (pp. 820-825). Hillsdale: LEA.

Fisher, Cynthia, Yael Gertner, Rose M. Scott, and Sylvia Yuan. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(2), 143-149.

Gahl, Susanne. (2008). *Time* and *thyme* are not homophones: The effect of lemma frequency on word duration in spontaneous speech. *Language*, *84*, 474-496.

Gerken, LouAnn & McIntosh, Bonnie J. (1993). Interplay of function morphemes and prosody in early language. *Developmental Psychology*, *29*(3), 448-457.

Hirsh-Pasek, Kathy, Kemler Nelson, Deborah G., Jusczyk, Peter W., Cassidy, Kimberly Wright, Druss, Benjamin, & Kennedy, Lori. (1987). Clauses are perceptual units for young infants. *Cognition*, *26*(3), 269-286.

Indefrey, P. & Levelt, P. J. W. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1-2), 101-144.

Johnson, Elizabeth K., & Jusczyk, Peter W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, *44*(4), 548-567.

Kemler Nelson, Deborah G., Hirsh-Pasek, Kathy, Jusczyk, Peter W., & Cassidy, Kimberly Wright. (1988). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*, *16*, 55-68.

Kisler, Thomas, Reichel, Uwe, & Schiel, Florian. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, *45*, 326–347.

Landau, Barbara, & Gleitman, Lila R. (1985). *Cognitive science series, 8. Language and experience: Evidence from the blind child.* Harvard University Press.

Lehiste, Ilse. (1970). Suprasegmentals. Cambridge: MIT Press.

Li, Aijun, Shi Rushen, & Hua, Wu. (2010), Prosodic cues to noun and verb categories in infant-directed Mandarin speech. *Speech Prosody*, *100088*, 1–4.

Lohmann, Arne, & Conwell, Erin. (2020). Phonetic effects of grammatical category: How category-specific prosodic phrasing and lexical frequency impact the duration of nouns and verbs. *Journal of Phonetics*, *78*, 1-22.

Maratsos, Michael P., & Chalkley, M. A. (1980). The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In Keith E. Nelson & Anne van Kleeck (Eds.), *Childrens Language, Volume 6* (pp. 127-214). Hillsdale: Lawrence Erlbaum Associates.

Mattys, Sven L., Jusczyk, Peter W., Luce, Paul A., & Morgan, James L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, *38*, 465-494.

Mintz, Toben H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, *90*(1), 91-117.

Moran, Steven, Blasi, Damian, Schikowski, Robert, Küntay, Aylin, Pfeiler, Barbara, & Stoll, Sabine (2018). A universal cue for grammatical categories in the input to children: Frequent frames. *Cognition*, *175*, 131-140.

Moran, Steven, & Cysouw, Michael. (2018). *The Unicode cookbook for linguists: Managing writing systems using orthography profiles.* Berlin: Language Science Press.

Moran, Steven, Schikowski, Robert, Pajović, Danica, Hysi, Cazim, & Stoll, Sabine. (2016). The ACQDIV Database: Min(d)ing the Ambient Language. In P*roceedings of the Tenth International Conference on Language Resources and Evaluation* (LREC 2016).

Naigles, Letitia G., Gleitman, Henry, & Gleitman, Lila, R. (1993). Children acquire word meaning components from syntactic evidence. In Esther Dromi (Ed.), *Language and cognition: A developmental perspective* (pp. 104-140). Ablex: Norwood.

Naigles, Letitia R., & Swensen, Lauren D. (2007). Syntactic supports for word learning. In Erika Hoff & Marilyn Shatz (Eds.), *Blackwell handbook of language development* (p. 212–231). Oxford: Blackwell.

Nazzi, Thierry, Betonocini, Josiane, & Mehler, Jacques. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 756-766.

Pinker, Steven (1984). *Language learnability and language development*. Cambridge: Harvard University Press.

Reddington, M., Chater, Nick, & Finch, Steven. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, *22*(4), 425-469.

Rochester, S. R. (1973). The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research*, *2*, 51-81.

Saffran, Jenny R., Aslin, Richard N., & Newport, Elissa L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-1928.

Schikowski, Robert, Paudyal, Netra P., & Bickel, Balthasar. (2011). Flexible valency in Chintang. In Bernard Comrie & Andrej Malchukov (Eds.), *Valency classes in the world's languages* (pp. 669-707). Berlin: De Gruyter Mouton.

Seifart, Frank, Strunk, Jan, Danielson, Swintha, Hatmann, Iren, Pakendorf, Brigitte, Wichmann, Søren, Witzlack-Makarevich, Alena, de Jong, Nivja H., & Bickel, Balthasar. (2018). Nouns slow down speech across structurally and culturally diverse languages. Proceedings of the National Academy of Sciences, 115, 5720-5725.

Shattuck-Hufnagel, Stefanie. & Turk, Alice. (2007). Multiple targets of phrase-final lengthening in American English words. Journal of Phonetics, 35(4), 445-472

Shi, Rushen, & Moisan, Annick. (2008). Prosodic cues to noun and verb categories in infant-directed speech. In Proceedings of the 32th annual Boston University conference on language development (pp. 450-461). Boston: Cascadilla Press.

Sorenson, John M., Cooper, William E., & Paccia, Jeanne M. (1978). Speech timing of grammatical categories. Cognition, 6(2), 135-153.

Stoll, Sabine, Lieven, Elena, Banjade, Goma, Bhatta, Toya Nath, Gaenszle, Martin, Paudyal, Netra P., Rai, Manoj, Rai, Novel Kishor, Rai, Ichchha P., Zakharko, Taras, Schikowski, Robert, & Bickel, Balthasar. (2015). Audiovisual corpus on the acquisition of Chintang by six children.

Turk, Alice, & Shattuck-Hufnagel, Stefanie. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, *35*, 445– 472.

Wasow, Thomas. (1997). End-weight from the speaker's perspective. *Journal of Psycholinguistic Research*, *26*(3), 347-361.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. The Journal of the Acoustical Society of America, 91(3), 1707-1717.

Proceedings of the 44th annual Boston University Conference on Language Development

edited by Megan M. Brown
and Alexandra Kohut

Cascadilla Press   Somerville, MA   2020

Copyright information

ISSN 1080-692X
ISBN 978-1-57473-057-9 (2 volume set, paperback)

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com