Title
Why do children pay more attention to grammatical morphemes at the ends of sentences?

Journal
Journal of Child Language, 45(3)

ISSN
0305-0009

Author
SUNDARA, Megha

Publication Date
2018-05-01

DOI
10.1017/s0305000917000356

Peer reviewed

Why do children pay more attention to grammatical morphemes at the ends of sentences?

Journal:	*Journal of Child Language*
Manuscript ID	JCL-07-16-085.R1
Manuscript Type:	Brief Research Report
Keywords:	morphemes, sentence position, english, memory, prosody

Why do children pay more attention to grammatical morphemes at the ends of sentences?

Running title: recency or final lengthening

Address for correspondence:

Megha Sundara

Department of Linguistics

University of California at Los Angeles

3125 Campbell Hall

Los Angeles, CA 90095-1543

Email: megha.sundara@humnet.ucla.edu

Word count – 3851

Abstract

Children pay more attention to the beginnings and ends of sentences rather than the middle. In natural speech, ends of sentences are prosodically and segmentally enhanced; they are also privileged by sensory and recall advantages. We contrasted whether acoustic enhancement or sensory and recall-related advantages are necessary and sufficient for the salience of grammatical morphemes at the ends of sentences. We measured 22-month-olds' listening times to grammatical and ungrammatical sentences with 3rd person singular *–s*. Crucially, by cross-splicing the speech stimuli, acoustic enhancement and sensory and recall advantages were fully crossed. Only children presented with the verb in sentence-final position, a position with sensory and recall advantages, distinguished between the grammatical and ungrammatical sentences. Thus, sensory and recall advantages alone were necessary and sufficient to make grammatical morphemes at ends of sentences salient. These general processing constraints privilege ends of sentences over middles, regardless of the acoustic enhancement.

Introduction

Edges of utterances, beginnings and particularly ends, are salient for children acquiring language (Slobin, 1973; 1985). Several influential theories, particularly Newport's "less is more" hypothesis (1990; Elman, 1993), propose that processing limitations privilege edges of utterance and might confer advantages to the beginning language learner. In this paper we investigate the roots of these processing limitations.

The processing advantage for speech material adjacent to edges of utterances is observed very early in life. Neonates (Ferry, Fló, Brusini, Cattarossi, Macagno, Nespor & Mehler, 2016) as well as 7-month-olds (Benavides-Varela & Mehler, 2015) detect switched syllables at edges but not in the middle of multisyllabic sequences. Infants also segment words earlier in development when they are presented at edges of utterances than in the middle (Seidl & Johnson, 2006; Seidl & Johnson, 2008) and associate such words with visual referents (Shukla, White & Aslin, 2011).

In fact, across languages, mothers typically place novel words at the ends of multiword utterances, even when the resulting sentences are ungrammatical (Aslin, Woodward, LaMendola, & Bever, 1996). This early processing advantage for speech material aligned with utterance edges continues into adulthood. Adults learn novel words in a non-native language more easily when the words are utterance-final than when they are utterance-medial (Golinkoff & Alioto, 1995). Enhanced encoding of material located at the edges of sequences has also been demonstrated in artificial language experiments (Endress, Nespor & Mehler, 2009).

Some researchers have suggested that the different rates at which various inflectional morphemes are acquired could relate to their distribution at the edges of sentences (Hsieh, Leonard & Swanson, 1999). Hsieh et al., examined conversation and stories addressed to children. They report that 52% of nouns with plural *–s* occur in sentence-final position, whereas only 16% of verbs with 3rd person singular *–s* occur sentence-finally. This results in the average duration of the 3rd person singular *–s* being about 25% shorter than that of the plural *–s,* due to the fact that utterance-final syllables are lengthened in English. Hsieh et al., (1999) suggest that this difference in distribution at edges, resulting in a difference in duration might account for English-learning children's earlier acquisition of plural *–s* compared to the 3rd person singular *–s*.

Although it is difficult to compare acquisition rates across different inflectional morphemes due to confounding differences in grammatical complexity, a comparison of children's production of the same morpheme across different sentence positions also shows earlier acquisition of the morpheme at the ends of sentences. Longitudinal data from 1-3-year-olds and cross-sectional data from 2-year-olds (Song, Sundara, & Demuth, 2009; Sundara, Demuth & Kuhl, 2011) show that children produce 3rd person singular *–s* more accurately sentence-finally compared to sentence-medially, even when mean length of utterance (MLU), utterance length, and final consonant (coda) complexity of the inflected verb is controlled. A similar pattern of results is also seen in production data on finite verb morphology – tense and agreement – in typically developing children as well as children with language impairment (Dalal & Loeb, 2005; cf. Leonard, Miller & Owen, 2000).

The present study

What makes speech material at the edges of sentences or utterances salient? Speech material at the edges of utterances may be salient due to *acoustic enhancement*. Languages use prosodic and segmental modifications to mark the edges of utterances. The edges of utterances in speech to children and adults are typically marked prosodically with intonation contours and pause duration (Beckman & Pierrehumbert, 1986; Fernald & Mazzie, 1991; Fisher & Tokura, 1996). They also involve segmental modifications such as initial strengthening or final lengthening (Bernstein Ratner, 1986; Fisher & Tokura, 1996; Fougeron & Keating, 1997; Horne, Strangert, & Heldner, 1995; Keating, Cho, Fougeron, & Hsu, 2003; Wightman, Shattuck-Hufnagel, Osterdorf, & Price, 1992).

Alternatively, speech material at the edges of sentences may also be salient because of *sensory or recall advantages* relating to its position. Here, we are interested primarily in the salience of ends, so that is what we discuss. First, material that follows can significantly reduce the audibility of auditory stimuli in utterance-medial positions (Moore, 1997). Whether due to the immaturity of the sensory or neural systems, or due to non-sensory factors like attention and memory, these effects of backward masking are more detrimental to the performance of children than adults (Saffran, Werker, & Werner, 2006; Hartley, Wright, Hogan, & Moore, 2000). Because the ends of utterances are not followed closely by other speech material, the audibility of segments at the ends of utterances are less likely to affected by backward masking. This tends to make speech material at the ends of sentences more salient compared to speech material in sentence-medial position. Second, in recall studies, the first (primacy) and last items (recency) are routinely remembered more accurately and more often (Deese & Kaufman, 1957); enhanced memory for the first and last items is observed for linguistic, visual and spatial stimuli (Gupta, Lipinski, Abbs & Lin, 2005; Hurlstone, Hitch & Baddeley, 2014)

In this study, we investigated whether the acoustic enhancement at the edges of utterances is necessary for the salience of material at edges. We tested the ability of 22-month-olds to detect the presence vs. absence of the 3rd person singular *–s* when the verb was embedded in a 3-word sequence, sentence-medially vs. sentence-finally. Previously, in Sundara et al (2011), we have shown that 22-month-olds listen significantly longer to grammatical compared to ungrammatical sentences when the 3rd person singular –*s* is in sentence-final position. Older children, that is, 27-month-olds as well detect the presence vs. absence of the 3rd person singular –*s* in sentence-final position, but by listening significantly longer to the ungrammatical sentences. In contrast, children at 22- and 27-months listen comparably to the presence vs. absence of 3rd person singular –*s* in sentence-medial position*.* Thus, children successfully detect the presence versus absence of 3rd person singular –*s* only in sentence-final position.

The children in Sundara et al., were tested on naturally produced sentences. That is, all sentence-final verbs were also acoustically enhanced, as is typical given final lengthening in English. Conversely, the sentence-medial verbs were not enhanced. Here we compare the performance of 22-month-olds on naturally occurring (a) final-enhanced, and (b) medial conditions previously presented in Sundara et al (control conditions), to two new experimental conditions. In the experimental conditions, all stimuli were cross-spliced. Specifically, we cross-spliced the acoustic instantiation of the inflected verb from one sentence position to the other. That is, in the medial cross-spliced condition, the verb from a sentence such as "*There he cries*" was excised, and replaced the verb in the sentence "*He cries now*". Similarly, in the final cross-spliced condition, the sentence-medial verb from "*He cries now*" replaced the verb in "*There he cries*". This cross-splicing was also applied to ungrammatical counterparts (e.g. "*There he cry*"). With this stimulus manipulation, verbs in the medial cross-spliced condition were acoustically enhanced but did not have sensory or recall benefits. In contrast, in the final cross-spliced condition no acoustic enhancement was available; however, the sensory and recall benefits were still available due to the placement of the verb sentence-finally.

If acoustic enhancement is necessary and sufficient to make grammatical inflections salient, in the current experiment children were expected to succeed only in the medial cross-spliced condition. Alternately, if sensory and recall advantages alone are necessary and sufficient to make grammatical inflections salient, children were expected to succeed only in the final cross-spliced condition.

Methods

Participants

All participants were full-term, monolingual English-learning children. According to parental report, the children had normal hearing, vision, and good health; none of the children had a cold or an ear infection on the day of testing. The final sample included 34 22-month-olds (20 girls, 14 boys, Range: 647 - 693 days) in the control group, tested using natural sentences (previously reported in Sundara et al., 2011), and 34 22-month-olds (16 girls, 18 boys, Range: 654 - 695 days) in the experimental group, tested using the cross-spliced stimuli. An additional nine children were tested but excluded from analysis either because they never looked away from the screen (2), did not complete testing (3) or due to experimenter error (4).

Table 1. Characteristics of the target verbs: MacArthur CDI comprehension and production scores for each target verb at 16-months (Dale & Fenson, 1996), as well as information from the CHILDES database regarding inflected with 3rd person singular *–s* and non-inflected verb frequency in child-directed speech (Li & Shirai, 2000; MacWhinney, 2000) are presented.

	Proportion of children from CDI database		Frequency from the CHILDES database	
Target verb	Comprehending at 16-months	Producing at 16-months	Inflected	Non-inflected
Cry	63.9	19.4	38	296
Throw	77.8	9.7	24	858

Eat	84.7	19.4	135	3960
Sleep	61.1	15.3	56	822

Stimuli

The stimuli were the same as the ones used in the perception experiment in Sundara et al., (2011). Highly frequent, familiar and pictureable verbs (Table 1) were embedded in 3-syllable, 3-word sentences with a 3rd person singular subject (e.g., *He cries now, There he cries*). The sentences are listed in Table 2. Each sentence was paired with an animated cartoon depicting the action (Figure 1; for details on the validation of the cartoon-sentence pair, see Sundara et al., 2011).

Figure 1. Animated cartoons paired with the verbs – cry, throw, eat and sleep.

In an animated voice, a 36-year-old, female native speaker of American English who is also a trained musician read 16 grammatical and ungrammatical sentences, four pairs where the verb was in sentence-final position and four pairs where the verb was in sentence-medial position. Sentences were recorded in a soundproof booth using a Shure SM81 table-top microphone. In the ungrammatical sentences, the 3rd person singular –*s* was omitted (e.g., *He* *sleeps now* vs. *He sleep now*). All sentences were digitized at a sampling frequency of 44.1 KHz and 16-bit quantization, and were excised using PRAAT (Boersma & Weenink, 2005).

To confirm that the grammatical and ungrammatical sentences did not differ systematically on extraneous prosody, the stimuli were low-pass filtered and presented to five native English-speaking adults (Mean age = 19.4; Range = 19:20). Low-pass filtering eliminates most segmental information, particularly the presence or absence of the 3rd person singular *–s,* while retaining rhythm, intonation and phrasing differences. Listeners were at chance when asked to determine whether the sentences were grammatical or ungrammatical (Mean percent correct = 52.5%; SD = 1.9). Thus, grammatical and ungrammatical sentences could not be distinguished based on their prosody alone.

Crucially, for the experimental group, the verbs in final and medial position were cross-spliced. Specifically, the verbs recorded in sentence-medial position were excised and replaced the verbs in sentence-final position and vice versa. Stimuli were cross-spliced to preserve all acoustic enhancements associated with utterance edges. These formed the stimuli for the final and the medial cross-spliced condition respectively. Thus, between the control and the experimental group acoustic enhancement and sentence position effects were fully crossed. In the control group, the sentence-final condition also naturally had acoustically enhanced verbs, whereas the medial condition did not. In the experimental group, the acoustically enhanced verbs were in the sentence-medial position and not in the sentence-final position.

Next, to ensure that final and medial cross-spliced stimuli did not differ in naturalness, ratings from five native English-speaking, To-BI trained adults were obtained. Listeners were asked to rate the naturalness of cross-spliced sentences on a scale from 1-7 (where 1 = completely unnatural, and 7 = completely natural). The ratings for final cross-spliced sentences

(Mean = 5.6; Median = 6; Range = 2:8) and medial cross-spliced sentences (Mean = 6.15; Median = 6; Range = 3:9) were entirely overlapping.

Table 2. The duration of 3rd person singular –*s* and the preceding vowel as well as the pitch excursion on the verb in each target sentence in cross-spliced stimuli is presented. The sentence-medial verbs from natural utterances were cross-spliced into sentence-final position, and vice versa. Thus, contradictory to acoustics of spoken English, the duration of the 3rd person singular –*s* morpheme as well as the vowel preceding it are longer in the medial cross-spliced condition. Note that for children tested on the control group with natural stimuli, the durations and pitch excursions for the final and medial condition are reversed.

Cross-spliced Condition	Sentence	Durations (*ms*)		Pitch excursion over the verb (*Hz*)	Intensity (dB)
		Preceding vowel	3rd person singular -*s*		
Final	There he cries	368	97	111	77.3
	There he throws	249	134	108	75.0
	Here she eats	179	117	67	73.9
	There he sleeps	144	146	138	74.1
Medial	He cries now	570	200	67	74.0
	He throws	459	228	280	75.2

fast				
She eats now	250	234	267	76.3
He sleeps now	238	215	126	70.2

The duration of the 3rd person singular –*s* and the vowel preceding it are presented in Table 2. Recall that, due to cross-splicing both duration measures are expected to be longer in the medial cross-spliced condition compared to the final cross-spliced condition. As expected, 3rd person singular –*s* was significantly longer in the medial cross-spliced condition (M = 219 ms; SD = 15) compared to the final cross-spliced condition (M = 124 ms; SD = 21), $t(3) = 9.49$, $p = 0.002$. Similarly, the preceding vowel duration was also significantly longer in the medial cross-spliced condition (M = 379 ms; SD = 162) compared to the final cross-spliced condition (M = 235 ms; SD = 99), $t(3) = 4.0$, $p = 0.03$.

Besides signaling utterance edges, longer segmental durations in English may also signal new information, particularly when accompanied by increased intensity and greater pitch excursions (Breen, Fedorenko, Wagner & Gibson, 2010; Pierrehumbert & Hirschberg, 1990). The extent of pitch excursion in the medial (M = 185 Hz; SD = 105) and final (M = 106 Hz; SD = 29) cross-spliced condition in our stimuli was comparable, $t(3) = 1.27$, $p = 0.29$. And so was the intensity [medial (M = 75dB; SD = 1.6); final (M = 74dB; SD = 2.7); $t(3) = 1.27$, $p = 0.5$]. Thus, the cross-splicing did not inadvertently signal a focus condition in one or other sentence position, although the longer duration in the medial, enhanced condition is certainly consistent with an interpretation of focused, new information.

Procedure

During testing, children sat on their parent's lap in a dark room facing a TV monitor. Audio stimuli were played at a comfortable 77 dB SPL over Bose loudspeakers placed next to the TV monitor, behind a dark curtain. The parent as well as the tester listened to music over sound attenuating JTC Clearwater headphones so as not to influence the child's behavior. A Sony SuperExWave camera lens was placed below the monitor. A tester outside the room was able to record the infant's gaze by watching the infant over a second TV monitor connected to the camera.

Children were tested using a modified version of the central fixation auditory preference procedure (Pinto, Fernald, McRoberts, & Cole, 1999) as in Sundara et al., (2011). The procedure was infant-controlled and implemented using Habit X (Cohen, Atkinson, & Chaput, 2004). At the beginning of each trial, a red flashing light accompanied by a baby giggle drew the child's attention to the screen. Once the child's gaze was on the screen, an animated cartoon was presented for the duration of the trial or until the child looked away from the TV screen for more than 2 seconds.

Testing was done in three phases. In the familiarization phase, children were presented with the 4 animated cartoons, each representing one verb, one-by-one, with no audio signal (Maximum trial duration = 10 seconds). The order of presentation of the 4 cartoons was randomized across children. A video-only, familiarization phase was necessary because children found the cartoons very interesting and would otherwise never look away from the screen during the test phase. In the test phase, children were presented with two blocks of 8 trials each (16 trials). On each trial, children saw one cartoon accompanied by a grammatical or ungrammatical sentence presented repeatedly (Maximum trial duration = 18.5 seconds). The order of presentation of the 4 grammatical and 4 ungrammatical sentences was randomized in each block. An experimenter, who was blind to the condition, coded how long the child looked at the monitor to obtain a measure of listening time to grammatical and ungrammatical sentences. Finally, to ensure that children tested in the experimental group were paying attention, one final post-test trial was presented. In the post-test trials infants saw the cartoon of a boy throwing a baseball, while listening to "He sleeps now". The listening time to the post-trial for all children was close to ceiling (17-18.5 seconds) confirming that they were surprised by the incongruence between the cartoon and the sentence presented auditorily. Thus, children were paying attention to the congruence of the audio and video stimuli; data from the post-test trial are not analyzed further. Children took about 10 minutes to complete this perception task.

Results

Listening time was analyzed in a 4-way ANOVA with Block (2 levels, 1~2) and Trial-type (2 levels, grammatical~ ungrammatical) as within-subjects variables and Sentence Position (2 levels: medial~final) and Acoustic Enhancement (yes~no) as between-subjects variables (SPSS version 24). Unsurprisingly, the main effect of block was significant, $F(1, 64) = 52.6$, $p < 0.00$, indicating that children's attention to the stimuli reduced over the blocks. A decline in listening times over successive blocks is a hallmark of infant listening time paradigms. In fact, listening times in Block 1 were nearly at ceiling – listening times to 75% of grammatical and ungrammatical were over 15s (where maximum trial duration is 18.5s). This was the case for both Experimental and Control conditions.

There was also a marginally significant effect of Acoustic Enhancement, $F(1, 64) = 4.1$, $p = 0.05$; and a significant interaction between Trial-type and Sentence Position $F(1, 64) = 8.1$, $p = 0.006$. Crucially, there was a significant 3-way interaction between Trial-type, Sentence Position and Acoustic Enhancement, $F(1, 64) = 6.7$, $p = 0.01$. Thus, 22-month-olds' listening time to grammatical and ungrammatical sentences varied as a function of Sentence Position and Acoustic Enhancement.

Next, we analyzed the results for the two sentence positions separately in a 3-way ANOVA. First we compared the two experiments where verbs were presented in medial position with Block (1~2) and Trial-type (grammatical~ungrammatical) as within-subjects factors and Acoustic enhancement (yes~no) as a between-subjects factor. Only the main effect of Block [$F(1, 32) = 36.2$, $p < 0.001$, $\eta_p^2 = 0.53$] was significant. There was only one other marginally significant interaction of Trial-type and Acoustic Enhancement [$F(1, 32) = 4.1$, $p = 0.05$, $\eta_p^2 = 0.11$]; all other effects were non-significant [p's > 0.2]. A follow-up paired comparison between listening time to grammatical and ungrammatical trials in Block 2, was in opposite directions for the enhanced and un-enhanced condition, though neither were significantly different from each other [enhanced: $t(16) = 1.2$, $p = 0.24$; not-enhanced: $t(16) = -1.7$, $p = 0.12$]. Unsurprisingly, there were no significant differences in Block 1 because looks were mostly at ceiling. Thus, 22-month-olds listened comparably to the grammatical and ungrammatical sentences when verbs were in sentence-medial position, regardless of acoustic enhancement.

Then, we compared the two experiments where the verbs were presented in final position using another 3-way ANOVA. Again, Block (1~2) and Trial-type (grammatical~ungrammatical) were within-subjects factors and Acoustic enhancement (yes~no) was as a between-subjects factor. As seen previously, the main effect of Block was significant [$F(1, 32) = 18.2$, $p < 0.001$, $\eta_p^2 = 0.36$]. Crucially, the main effect of Trial-type was also significant, $F(1, 32) = 12.6$, $p = 0.001$, $\eta_p^2 = 0.28$. A follow-up paired comparison between listening time to grammatical and ungrammatical trials in Block 2 was significant for both the acoustically enhanced [$t(16) = 2.5$, p

= 0.02], and not-enhanced condition [$t(16) = 2.4$, $p = 0.03$]. Note that like in the medial position, there were no significant differences in Block 1 in final position either because looks were mostly at ceiling. Thus, 22-month-olds listened significantly longer to grammatical compared to ungrammatical sentences when verbs were in sentence-final position, regardless of acoustic enhancement. In sum, 22-month-olds were able to detect the presence vs. absence of 3rd person singular –*s* regardless of acoustic enhancement in sentence-final position. Average listening times in Block 2 are presented in Figure 2, and summarized in Table 3.

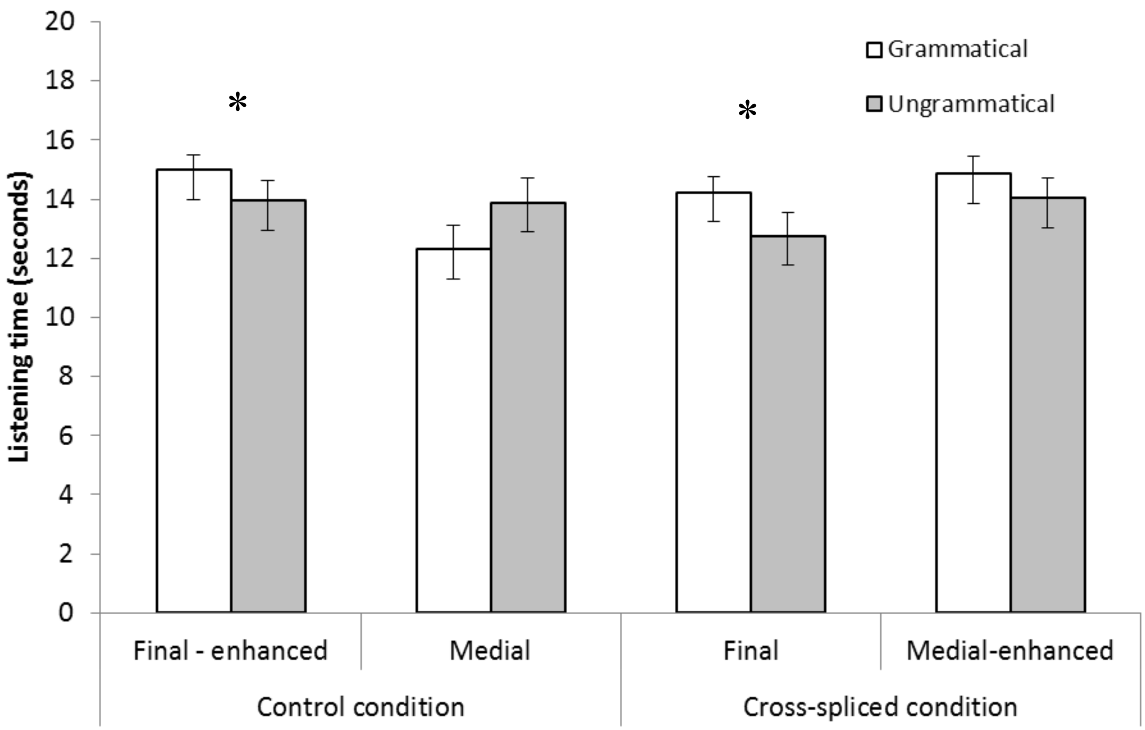

Figure 2. Average listening time (± *SE*) in Block 2 to grammatical and ungrammatical sentences in the control and cross-spliced groups is presented. The asterisks mark the conditions where the listening times to the grammatical and ungrammatical sentences were significantly different in a paired t-test.

Condition	Listening time (s)	
	Grammatical	Ungrammatical
Final-enhanced (Sundara et al., 2011)	15.0 (0.52)	13.9 (0.68)
Medial (Sundara et al., 2011)	12.3 (0.81)	13.9 (0.85)
Final Cross-spliced	14.2 (0.55)	12.8 (0.79)
Medial Cross-spliced	14.9 (0.57)	14.0 (0.67)

Table 3. Average listening times in seconds in Block 2 (SE) to grammatical and ungrammatical sentences

Discussion

We tested whether acoustic enhancement of morphological inflections occurring at the ends of sentences was necessary and sufficient to account for the salience of speech material at the ends of sentences. For this, we compared 22-month-olds' ability to detect the presence or absence of 3rd person singular *–s* in four conditions. In the control conditions, natural sentences with medial, non-enhanced morphemes were contrasted with final, enhanced morphemes (previously reported in Sundara et al., 2011). In the experimental conditions, the sentence stimuli were cross-spliced. In the medial cross-spliced condition, acoustic enhancement was preserved. In the final cross-spliced condition, only the sensory and recall advantages were preserved.

Our results show that 22-month-olds successfully detected the presence vs. absence of 3rd person singular *–s* in sentence-final position, regardless of acoustic enhancement. Thus, the sensory and recall advantages conferred by placement at the end of a sentence alone are necessary and sufficient for the salience of edge-aligned grammatical inflections.

That acoustic enhancement alone is not sufficient to make inflections placed sentence-medially salient is consistent with the results of an on online perception study by typically-developing children and children with SLI (Montgomery & Leonard, 2006). Acoustic manipulation of the duration of 3rd person singular and possessive *–s* in this study was primarily restricted to target words in non-final position and thus, did not successfully enhance the performance of either group of children.

There are cross-linguistic implications of our finding that sensory and recall advantages make grammatical inflections at the ends of sentences salient. Sensory and recall advantages being domain general likely confer advantages for learners of any language. Specifically, facilitative effects of edge-alignment are likely to be observed in morphological acquisition cross-linguistically, with children acquiring inflections earlier at the edges of utterances.

In conclusion, we show that infants' detection of inflectional morphemes – 3rd person singular *–s* in this case, is facilitated when the verbs are placed at the edge of an utterance. Additionally, acoustic enhancement resulting segmental and prosodic manipulations alone was not necessary or sufficient to make grammatical inflections salient. However, it is possible that these acoustic enhancements add to the salience of speech material at utterance edges. Future investigations are needed to test the additive effects of acoustic enhancements, if any, on the salience of edge-aligned grammatical morphemes. Such an investigation could also serve to disentangle the relative contribution of segmental and prosodic enhancements.

Acknowledgments

References

Aslin, R. N., Woodward, J. Z., LaMendola, P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 117-134). Mahwah, NJ: Lawrence Erlbaum Associates.

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational Structure in Japanese and English. *Phonology Yearbook, 3*, 255-309.

Benavides-Varela, S. and Mehler, J. (2015), Verbal Positional Memory in 7-Month-Olds. *Child Development*, 86: 209–223.

Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother–child speech. *Journal of Phonetics, 14*, 303–309.

Boersma, P., & Weenink, D., (2005). Praat : Doing phonetics by computer. Software. University of Amsterdam.

Breen, M., Fedorenko, E., Wagner, M., Gibson, E., (2010). Acoustic correlates of information structure. *Language and Cognitive Processes,* 25 (7), 1044-1098.

Cohen, L. B., Atkinson, D. J., & Chaput, H. H. (2004). A new program for obtaining and organizing data in infant perception and cognition studies (Version 1.0). Austin: University of Texas.

Dalal, R., & Loeb, D. (2005). Imitative production of regular past tense *-ed* by English-speaking children with specific language impairment. *The International Journal of Communication Disorders, 40*(1), 67-82.

Dale, P.S. & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, 28, 125-127.

Deese, J., & Kaufman, R. A. (1957). Serial effects in recall of unorganized and sequentially organized verbal material *Journal of Experimental Psychology, 54*(3), 180-187.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48*, 71-99.

Endress, A.D., Nespor, M. & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*, 13(8), 348-353.

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology, 27*, 209–221.

Ferry, A. L., Fló, A., Brusini, P., Cattarossi, L., Macagno, F., Nespor, M., & Mehler, J. (2016). On the edge of language acquisition: inherent constraints on encoding multisyllabic sequences in the neonate brain. *Developmental science, 19(3),* 488-503.

Fisher, C., & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: cross-linguistic evidence. *Child Development, 67*, 3192–3218.

Fougeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of Acoustical Society of America, 101*, 3728–3740.

Golinkoff, R. M., & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: Implications for language acquisition. *Journal of Child Language, 22*, 703-726.

Gupta, P., Lipinski, J., Abbs, B., & Lin, P.-H. (2005). Serial position effects in nonword repetition. *Journal of Memory and Language*, 53(1), 141-162.

Hartley, D. E. H., Wright, B. A., Hogan, S. C., & Moore, D. R. (2000). Age related improvements in auditory backward and simultaneous masking in 6- to 10-year-old children. *Journal of Speech, Language, and Hearing Research, 43*, 1402-1415.

Horne, M., Strangert, E., & Heldner, M. (1995). Prosodic boundary strength in Swedish: Final lengthening and silent interval duration. In K. Elenius (Ed.), *Proceedings of the 13th International Congress of Phonetic Sciences* (Vol. 1, pp. 170-173). Stockholm, Sweden.

Hsieh, L., Leonard, L. B., & Swanson, L. A. (1999). Some differences between English plural noun inflections and third singular verb inflections in the input: The contributions of frequency, sentence position, and duration. *Journal of Child Language, 26*, 531-543

Hurlstone, M.J., Hitch, G.J., & Baddeley, A.D. (2014). Memory for serial order across domains: an overview of the literature and directions for future research. Psychological Bulletin, 140 (2), 339–373.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound pattern of words in fluent speech. *Cognitive Psychology, 23*, 1-23.

Keating, P., Cho, T., Fougeron, C., & Hsu, C.-S. (2003). Domain-initial articulatory strengthening in four languages. In J. Local, R. Ogden & R. Temple (Eds.), *Papers in Laboratory Phonology VI* (pp. 143–161). Cambridge: Cambridge University Press.

Leonard, L. B., Miller, C., & Owen, A. (2000). The comprehension of verb agreement morphology by English-speaking children with specific language impairment. *Clinical Linguistics & Phonetics, 14*, 465-481.

Li, P., & Shirai, Y. (2000). The acquisition of lexical and grammatical aspect. Berlin & New York: Mouton de Gruyter.

MacWhinney, B. (2000). The CHILDES Project: Tools for analyzing talk. Mahwah, NJ: Lawrence Erlbaum Associates.

Mattys, S. L., & Jusczyk, P. W. (2001). Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology: Human Perception and Performance, 27*(3), 644-655.

Montgomery, J. W., & Leonard, L. B. (2006). Effects of acoustic manipulation on the real-time inflectional processing of children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 49*, 1238-1256.

Moore, B. C. J. (1997). *An introduction to the psychology of hearing* (4th ed.). London: Academic Press.

Newport, E. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11-28.

Pinto, J. P., Fernald, A., McRoberts, G. W., & Cole, S. (1999). Reliability and validity in infant auditory preference procedures. In C. Rovee-Collier, L. Lipsitt & H. Hayne (Eds.), *Advances in Infancy Research* (Vol. 12, pp. 221-236). Stamford, CT: Ablex.

Saffran, J. R., Werker, J. F., & Werner, L. A. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In R. Siegler & D. Kuhn (Eds.), *Handbook of Child Development* (pp. 58-108). New York: Wiley.

Seidl, A., & Johnson, E. K. (2006). Infant word segmentation revisited: Edge alignment facilitates target extraction. *Developmental Science, 9*(6), 565-573.

Seidl, A., & Johnson, E. K. (2008). Boundary alignment enables 11-month-olds to segment vowel-initial words from speech. *Journal of Child Language, 35*, 1-24.

Shukla, M., White, K. S., and Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences*, 108, 6038-6043.

Slobin, D. (1973) Cognitive Prerequisites for the development of Grammar. In C. A. Ferguson

& D. I. Slobin (Eds.), Studies of Child Language Development, 175-209.

Slobin, D. I. (1985). Crosslinguistic Evidence for the language-making capacity. In D. I. Slobin (Ed.), The crosslinguistic study of language acquisition, vol 2, 1157-1256.

Song, J.Y., Sundara, M. & Demuth, K. (2009). Effects of phonology on children's production of English 3rd person singular *–s*. *Journal of Speech, Language and Hearing Research, 52(3),* 623-642.

Sundara, M., Demuth, K. & Kuhl, P.K. (2011). Sentence-position effects on children's comprehension & production of English 3rd person singular *–s*. *Journal of Speech, Language and Hearing Research, 54,* 55-71.

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