INVESTIGATING METRICAL CONTEXT EFFECTS ON ANTICIPATORY COARTICULATION IN CONNECTED SPEECH DEVELOPMENT

Authors: Jillian Adkins, B.A.[1], Christina Gildersleeve-Neumann, Ph.D.[1], Melissa Redford, Ph.D.[2]
Affiliation(s): [1]Portland State University, [2]University of Oregon
Email: redford@uoregon.edu

ABSTRACT

If rhythm acquisition is influenced by the development of articulatory timing, then metrical structure might be expected to condition this timing. This study tested this hypothesis by investigating anticipatory effects of an upcoming noun on the production of a preceding determiner, under the assumption that anticipatory coarticulation indexes chunking. Simple S-V-O sentences were elicited from 5-year-olds, 8-year-olds, and adults. The V was either monosyllabic *packed* or disyllabic *patted*. The O was a determiner phrase where nouns varied either in onset place-of-articulation (POA; *tack* vs. *cat*) or in their rhymes (*tack* vs. *toot*). Acoustic analyses of determiner schwa F1 and F2 showed no effect of verb on schwa coarticulation. Given other results, including an interaction between age group and POA, the findings suggest that the acquisition of articulatory timing is independent of metrical structure, even if this timing is related to speech rhythm acquisition.

Keywords: prosody, speech development, coarticulation

1. INTRODUCTION

Rhythm is a highly salient feature of speech and a marker of difference between typical and atypical speech. Moreover, atypical rhythm patterns have been shown to increase the perception of disorder [1, 2], which is correlated with negative social evaluations [3, 4]. For these reasons, it is important to understand what drives rhythm acquisition: the more complete our understanding, the more able we are to effectively and efficiently intervene to help remediate disordered rhythm.

Measurement-based research on the typical acquisition of English rhythm suggests protracted development [5, 6]; similar to disordered speech rhythms, school-aged children's speech is more equally timed than adults' speech [7, 8]. In particular, their vowel-to-vowel durations are more nearly equal compared to adults' speech [9, 10, 11]. This specific finding suggests that rhythm may be tied to the development of articulatory timing, which is also develops very slowly [8, 12]. Yet, most research on speech rhythm acquisition has focused on the acquisition of prosodic structures, especially metrical structure [e.g., 13, 14, 15, 16] and associated lexical stress patterns [e.g., 17, 18]. Here, we investigate whether articulatory timing in school-aged children's speech is conditioned by these factors or is independent of them.

In previous research, we investigated grammatical word reduction as a function of metrical structure [19]. The current study extends this work to investigate how grammatical words are chunked with adjacent content words, where chunking is indexed by anticipatory coarticulation. According to metrical theory, grammatical words are less prominent than content words. Thus, monosyllabic grammatical words are preferentially chunked with a preceding monosyllabic content word to form a trochaic foot. When this chunking is not possible because the preceding content word is itself a trochee, the grammatical word is unfooted and so chunked according to syntactic structure [20]; for example, it is chunked with a following noun if the grammatical word is a determiner. This chunking pattern predicts that anticipatory effects on grammatical word production will be stronger in cases where the word is unfooted than in cases where it is footed. Of course, the prosodic word that results from the unfooted case is also an iamb. Some prior work has shown that children's production of iamb is still immature well into the school-aged years [18]. Thus, it could be that anticipatory effects on grammatical word production will emerge only slowly over developmental time.

2. METHODS

2.1. Participants

Sixteen five-year-olds, 16 eight-year-olds, and 16 adults participated in this study, and data from seven 5-year-olds, seven 8-year-olds, and four adults have

been analyzed. Participants were recruited through word of mouth. All participants are native-English speaking with typically-developing speech and language, according to parent report and assessment for children, and self-report for adults. All passed a pure tone hearing screening at 25 dB presented at 1000, 2000, and 4000 Hz. To determine typical speech and language development, children completed the articulation subtest of the Diagnostic Evaluation of Articulation and Phonology (DEAP) as well as four Core Language Score (CLS) subtests from the Clinical Evaluation of Language Fundamentals - 5th Edition (CELF-5). Participants who received a scaled score below seven on the DEAP or on the CLS were excluded from the study.

2.2. Stimuli

To test for effects of metrical structure on articulatory timing in school-aged children's and adults' speech, we elicited sentences with different verbs, as in Gerken [8]. Two target verbs were used: one was monosyllabic (*packed*) and one was disyllabic (*patted*). These verbs were combined with nouns in simple stimulus sentences that had the following shape: "*Maddy packed/patted THE NOUN today.*" According to metrical theory, the determiner in the "packed" sentences should be chunked with the verb to create the prosodic word, "*packed the,*" which follows the preferred trochaic strong-weak stress pattern of English. The determiner in the "patted" will remain unfooted because the verb ends with a weak syllable. It is thus only in this context that "the" is chunked with the following noun to form a prosodic word. These differences in stress patterns and footing are predicted to influence degree of anticipatory coarticulation in "the" production.

We test for both place-of-articulation (POA) effects on schwa in "the" and for effects of the following stressed vowel on schwa in "the". In the POA condition, the 4 monosyllabic noun targets all had the same vowel (*dad, tack, gak, cat*), but half of the words had simple alveolar stop onsets and the other half had simple velar stop onsets. In the V-V condition, all 4 monosyllabic noun targets had single alveolar stop onsets (*dad, dude, tack, toot*), but half of the words had a low front vowel and the other half had a high back vowel. Two target words were the same across the 2 conditions, thus altogether the determiner was produced adjacent to just 6 target nouns in two metrical contexts.

2.3. Procedure

The target sentences were blocked by metrical context (i.e., "patted" or "packed") and elicited in random order with other filler sentences. The aim was to elicit 4 correct and fluent productions of each target sentence, and so each block was repeated 4 times over the course of a 3-hour study session, with assessment and additional experimental tasks interspersed across the session. Picture prompts were used to facilitate elicitation. In addition, the experimenter provided a model sentence, at a conversational rate, then asked the child to repeat the production. To encourage naturalistic productions, and therefore naturalistic chunking, the examiner monitored the participant's production for disfluencies, unusual rate, over-enunciation, and pauses. If any errors were present, the sentence was elicited an additional time at the end of the block. Participants' speech was audio-video recorded for later analysis.

2.4. Analysis

The audio was stripped from the audio-video recording for the analyses reported here. Sentence waveforms were then displayed using Praat [21] and a textgrid file of each elicitation was created with five tiers to identify the metrical context (i.e., verb), block/repetition number, and vowels for measurement. Figure 1 provides an example of tiered segmentation of a target sentence, *"Maddy patted the gak today."*

Figure 1: Vowel segmentation in the target sentence, *Maddy patted the gak today*.
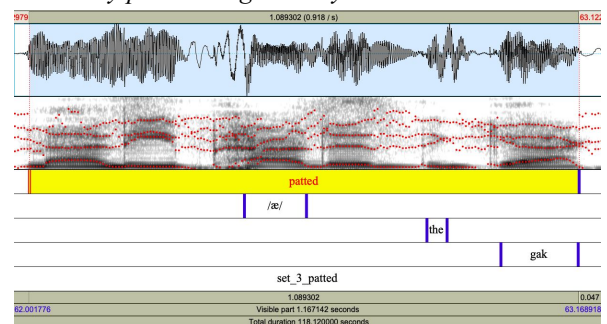
2.5. Measurements

Vowel F1 and F2 were taken at vowel midpoint. Although our focus is on the effects metrical context and of the following noun on "the" production, we measured F1 and F2 in both the schwa and in the target noun vowel. For the POA condition, F1 and

F2 were taken at noun vowel onset to confirm perseveratory effects of the consonants on the vowel. In total, 748 sentences were analyzed (5-year olds: 291, 8-year olds: 266, adults: 191); of those, 691 were acceptable for use (5-year olds: 261, 8-year olds: 248, adults: 182). Fifty-seven sentences were excluded due to ambiguous formant values.

2.6. Statistical Analysis

F1 and F2 measures were averaged across repetition within vowel (determiner or noun), POA (alveolar or velar) or target vowel (low front or high back), metrical context (footed or unfooted) and speaker. The effect of the within subjects factors were investigated using the mean values using repeated measures ANOVA with speaker's age group as a between subjects factor. Since formant values were not normalized for vocal tract length, a simple effect group was expected, especially on F2 values. This effect is not of particular interest in the context of the present study. Instead, we are interested in any interaction between group and the experimental factors. Interactions between group and POA or target vowel on "the" production would indicate articulatory timing differences across age groups. Interactions between group and metrical context on "the" production would indicate that metrical context conditioned articulatory timing differently across age groups.

3. RESULTS

### 3.1. Anticipatory V-to-C coarticulation

The analyses indicated no significant effects on the determiner vowel F1. It is possible that when more speakers are added to each age group, the effect of metrical context will be significant since the data trended towards significance [F(1,15) = 3.67, p = .075]. There were no other such trends in the data.

The results for determiner vowel F2 were more interesting. Although there was no simple effect of either metrical context or group, the effect of POA was significant [F(1,15) = 29.25, p < .001]. More importantly, there was a significant group by POA interaction [F(2,15) = 4.66, p = .027], indicating articulatory timing differences as a function of age. No other interactions were significant.

Figure 2 shows that the schwa in "the" was produced with a higher F2 before a velar consonant than before an alveolar consonant by speakers in across all age groups. The direction of this result is

consistent with an allophonic production of a palatalized /k/ before the front vowel. It is also consistent with the higher frequency burst typical of a /k/ release. The interaction between age group and POA was due to a stronger coarticulatory effect on adults' "the" production compared to children's "the" production.

Figure 2: The average F2 in *the* as a function of the speaker's age (5YO= 5-year-olds; 8YO - 8-year-olds; AD = adults) and the consonantal onset of the following monosyllabic noun.
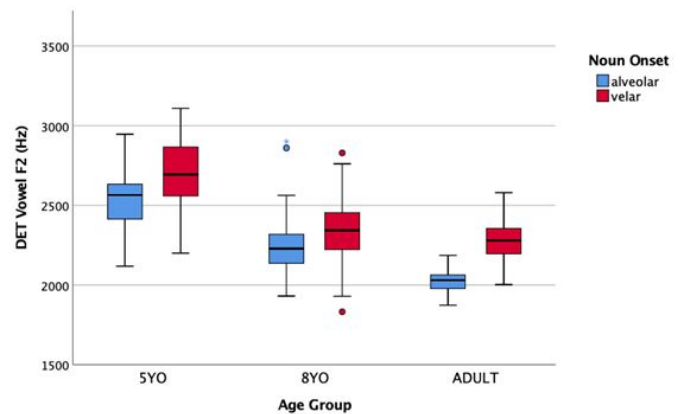

Stronger V-to-C coarticulation in adults compared to children might be due simply to more precise consonantal articulation in adults compared to children rather than to age-related differences in the chunking of grammatical words. To investigate this possibility, we tested for a group by POA interaction on F2 taken at the vowel onset in the noun (i.e., C-to-V coarticulation). The analyses revealed significant effects of group [F(2,15) = 24.017, p < .001] and POA [F(1, 15) = 48.637, p < .001], but no interaction between these factors. At least within a word, consonantal articulation appears to have the same effect on vowel articulation regardless of the speaker's age.

3.2. Anticipatory V-to-V coarticulation

In the target vowel condition, analyses showed unsurprising simple effects of age group on the determiner vowel F1 [F(2,15) = 3.394, p = .061] and F2 [F(2,15) = 24.108, p < .001]. Anticipatory effects of the upcoming stressed vowel on schwa were only observed for F1 [F(1,15) = 19.00, p = .001]. In particular, F1 was lower before target /u/ than before target /æ/. This effect did not vary with age group, as is evident from Figure 3..

Figure 3: The average F1 in *the* as a function of the speaker's age (5YO = 5-year-olds; 8YO = 8-year-olds; AD = adults) and the stressed vowel in the following monosyllabic noun.
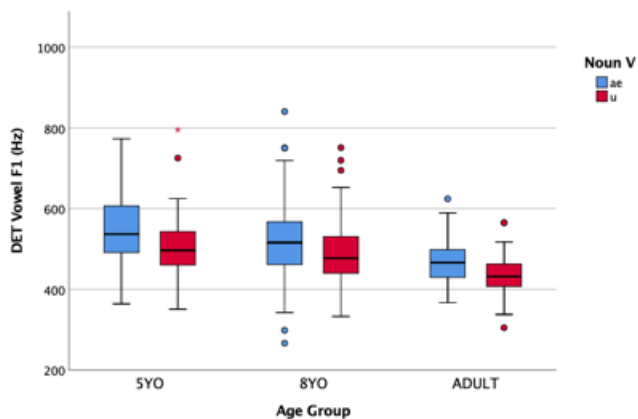
There was no effect of metrical context on "the" production in the target vowel condition. Thus, the V-to-V results suggest that the grammatical word may be chunked similarly regardless of metrical context and a speaker's age.

4. DISCUSSION AND FUTURE DIRECTIONS

If rhythm acquisition is influenced by the development of articulatory timing, then metrical structure might be expected to condition this timing. The current study tested this hypothesis by investigating anticipatory effects of an upcoming noun on the production of a preceding determiner, under the assumption that anticipatory coarticulation indexes chunking. Overall, the results suggest little to no effect of metrical structure on chunking in child and adult speech. The results so far also suggest little effect of age on coarticulatory effects of an upcoming noun on the production of the determiner: both children and adults showed similar anticipatory effects of the noun onset and noun vowel on the determiner vowel. The one exception was the significant group by POA interaction on determiner vowel F2: the effect of the noun onset was stronger in adults' speech than in children's speech. The finding could be due to the more rapid articulation rate of adults' speech compared to children's speech or to adults more precise segmental articulation or, perhaps, to differences in how children and adults chunk determiners with an adjacent content word. We will explore these different possibilities in future work with greater numbers of participants per group.

The type of research we report here has implications for how we understand the speech of children with speech and language disorders, especially when we can compare typical development to atypical development. Accordingly, we are beginning to explore metrical timing and articulatory constraints in children with known prosodic deficits: childhood apraxia of speech and autism. This research, and the understanding it affords, will inform better assessment and effective intervention for disorders characterized by atypical prosodic patterns.

6. REFERENCES

[1] Olejarczuk, P., Redford, M. A. (2013). The relative contribution of rhythm, intonation and lexical information to the perception of prosodic disorder. *Proceedings of Meetings on Acoustics ICA 2013* (Vol. 19, No. 1, p. 060154). ASA.

[2] Paul, R., Shriberg, L. D., McSweeny, J., Cicchetti, D., Klin, A., Volkmar, F. (2005). Brief report: Relations between prosodic performance and communication and socialization ratings in high functioning speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 35, 861.

[3] McCabe, P. C., Meller, P. J. (2004). The relationship between language and social competence: How language impairment affects social growth. *Psychology in the Schools*, 41, 313-321.

[4] Redford, M. A., Kapatsinski, V., & Cornell-Fabiano, J. (2018). Lay listener classification and evaluation of typical and atypical children's speech. *Language and Speech*, 61, 277-302.

[5] Sirsa, H., Redford, M. A. (2011). Towards understanding the protracted acquisition of English rhythm. *Proceedings of the International Congress of Phonetic Sciences. International Congress of Phonetic Sciences* (Vol. 2011, pp. 1862-1865). NIH Public Access.

[6] Polyanskaya, L., & Ordin, M. (2015). Acquisition of speech rhythm in first language. *Journal of the Acoustical Society of America*, 138, EL199-EL204.

[7] Allen, G., Hawkins. S. (1978). The development of phonological rhythm. In A. Bell & J. Bybee Hooper (eds.), Syllables and Segments (pp. 173–185). New York: North-Holland Publishing.

[8] Hawkins, S. (1984). On the development of motor control in speech: Evidence from studies of temporal coordination. In N. Lass (ed.) *Speech and Language* (Vol 11, pp. 317-374). Academic Press.

[9] Grabe, E., Post, B., Watson, I. (1999). The acquisition of rhythmic patterns in English and French. *Proceedings of the International Congress of Phonetic Sciences. International Congress of Phonetic Sciences* (Vol. 1999, pp. 1201-1204]

[10] Bunta, F., Ingram, D. (2007). The acquisition of speech rhythm by bilingual Spanish-and English-speaking 4-and 5-year-old children. *Journal of Speech, Language, and Hearing Research, 50,* 999-1014.

[11] Payne, E., Post, B., Astruc, L., Prieto, P., & Vanrell, M. D. M. (2012). Measuring child rhythm. *Language and Speech*, 55(2), 203-229.

[12] Redford, M.A. (2015). The acquisition of temporal patterns. In Redford, M.A. (ed.), *The handbook of speech production* (pp. 379-403). Boston: Wiley-Blackwell.

[13] Fikkert, P. (1994). On the acquisition of prosodic structure. PhD thesis, University of Leiden, The Netherlands.

[14] Gerken, L. (1996). Prosodic structure in young children's language production. *Language*, 72, 683-712.

[15] Kehoe, M., & Stoel-Gammon, C. (1997). Truncation patterns in English-speaking children's word productions. *Journal of Speech, Language, and Hearing Research*, 40(3), 526-541.

[16] Goffman, L., Malin, C. (1999). Metrical effects on speech movements in children and adults. *Journal of Speech, Language, and Hearing Research, 42*, 1003-1015.

[17] Kehoe, M., Stoel-Gammon, C., & Buder, E. H. (1995). Acoustic correlates of stress in young children's speech. *Journal of Speech, Language, and Hearing Research*, 38(2), 338-350.

[18] Ballard, K. J., Djaja, D., Arciuli, J., James, D. G., van Doorn, J. (2012). Developmental trajectory for production of prosody: Lexical stress contrastivity in children ages 3 to 7 years and in adults. *Journal of Speech, Language, and Hearing Research, 55*, 1822-1835.

[19] Redford, M. A. (2018). Grammatical word production across metrical contexts in school-aged children's and adults' speech. *Journal of Speech, Language, and Hearing Research,* 1-16.

[20] Selkirk, E. (1996). The prosodic structure of function words. In J.L. Morgan & K. Demuth (eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, (pp. 187-214). Mahwah, NJ: Erlbaum.

[21] Boersma, P., Weenink, D.. Praat: doing phonetics by computer [Computer program]. Version 6.0.35, retrieved March 2017 from http://www.praat.org/