The development of syllable timing by children with cochlear implants and their peers with normal hearing.

Mark Gibson[1], Ferenc Bunta[2], Charles Johnson[3] & Miriam Huárriz[1]
1. Universidad de Navarra 2. University of Houston 3. Brigham Young University

This research addresses the development of syllable timing by bilingual Spanish- and English-speaking children with cochlear implants (*henceforth*, CI group) and their peers with normal hearing (*henceforth*, NH group). Our variables were: voice onset time (*henceforth*, VOT), consonant duration in C and CC onsets, and vowel duration. Pre-lingually deafened children with cochlear implants were chosen because they did not have access to speech until after implantation at age 1, thus they did not have access to early word forms for the first year of life, which is when early word learning and perceptual attunement occur (Bergelson & Swingley, 2012; Vihman, 2017d). Bilinguals were chosen in order to study the impact of language interference/interaction, and to address the issue as to whether bilinguals access two different timing systems when producing speech, or apply the same timing patterns to two different languages (or alternately construe a hybrid timing system). Groups were matched on hearing age, where the average was 5;3 (five years and three months). We chose this age since calibrating speech forms using auditory feedback does not ensue until, minimally, age 4 (MacDonald et al., 2012; Ménardet al., 2008).

A word list was developed that contains over 80 real words to elicit the target items. The corpus was designed in order to obtain a range of possible phonological targets in word initial and word final positions. We opted to maximize lexical frequency at the expense of not having an optimally balanced corpus predominantly due to work by Edwards, Beckman & Munson (2007), who show that temporal cues are susceptible to word frequency effects in children (though we remedy this in the statistical models).

For the study, audio signals for 22 children (11 CI group and 11 NH group) were obtained at a sampling frequency of 44,100. Voiced (/b, g/) and voiceless (/p, t, k/) stops were targeted in word initial, singleton onsets and complex clusters (with /l/ and /ɹ/ as C2) using the same words for all participants to ensure consistency across the participants' samples and control for the phonetic environment of the test items.

For the English production data, the effects of the phonological specifications of interest (C1 place/voice, and complex/simplex onset) on the temporal variables were analyzed using the "lme4" package in R Studio. In all cases, models were tested for main effects and all interactions. Speaker and item were modeled as random effects. Random intercepts by subject and random by-subject slopes were included.

To compare the English productions and the same participant's Spanish productions reported in Gibson et al. (2018) Random Forest models in an unsupervised learning environment were used employing the scikit- learn package in Python. Each of our Random Forest classifiers consisted of 100 Decision Trees, each of which had a maximum depth of 2 (only two Boolean conditions maximum were allowed to classify any piece of data). Five experiments were created to compare 1) English productions between NH and CI groups, 2) Spanish productions between NH and CI groups, 3) English productions of NH and CI groups with adult native English speakers, 4) Spanish productions of NH and CI groups with adult native Spanish speakers, and 5) English and Spanish (mixed) productions with the CI and NH groups. The models were trained on data from monolingual English- and Spanish-speaking adults.

Finally, Pearson product moment correlation models were used to examine the relationship between speech rate and phonological knowledge on the participants' productions.

For VOT, the NH group showed a main effect based on onset complexity (χ2[1, N=199] = 4.05, p=0.044), though this effect did not obtain for the CI group (χ2[1, N=201] = 0.241, p=0.623). For the NH group, VOT lengthened as consonants were added to the onset. For the CI group, VOT only increased in CC for velar and coronal initial clusters (but not for labial initial clusters).

As for vowel duration, results for the NH group show that vowel duration is predicted in large part by onset complexity such that vowels extend temporally following complex clusters as compared to vowels following singleton word initial stops (χ2[1, N=206] = 10.59, p=0.001). However, for the CI group there was no effect of onset complexity on vowel duration (χ2[1, N=195] = 2.17, p=0.14), though a place*onset type interaction obtained (χ2[2, N=195] = 20.9, p=0.002).

Regarding lateral duration in C2 and in word initial singleton position, as predicted, and in line with the children's Spanish productions reported in Gibson et al. (2018), no effect of group was found (χ2[2, N=426] = 2.23, p=0.14). Both groups produced the compression effect for laterals that typifies English (and Spanish).

Our Random Forest analyses (which will be presented in full) were able to categorize with substantial accuracy the children's productions in each language, suggesting that the children do indeed access separate temporal plans when producing each language. Whether or not they are accessing two gradient dimensions of the same phonological system as per Antoniou et al. (2010), or accessing two distinct phonological systems remains an open empirical question we intend to examine in moving forward.

Our Pearson product moment regressions showed strong correlations between vowel duration and speech rate for both groups in English (CI: [r = 0.22, n =321, p = 0.003]; NH: [r = 0.53, n =279, p < 0.001]) and Spanish (CI: [r = 0.56, n =299, p < 0.001]; NH: [r = 0.60, n =315, p < 0.001]). Vowels lengthen as speech rate reduces (after complex onsets). As for the effects of phonological knowledge on the speech production tasks, results show a strong correlation between phonological awareness and lateral stability for Spanish ([r = -0.71, n =22, p < 0.001]), but not English ([r = 0.11, n =44, p = 0.64]) , which we surmise is due to the fact that Spanish only has one clear lateral /l/ for all syllabic positions, while English has two (/l/ and /ɫ/).

References:

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9), 3253–3258.

Beckman, M., Munson, B. & Edwards, J. (2007). The influence of vocabulary growth on developmental changes in types of phonological knowledge. In: Cole J, Hualde J, editors. *Laboratory Phonology*. Vol. 9. New York: Mouton de Gruyter, 241–264.

Gibson, M., Bunta, F., Goodin-Mayeda, E., Hernandez, A. (2018). The acquisition of syllable-level timing contrasts by English-and Spanish-speaking bilingual children with normal hearing and English-and Spanish-speaking bilingual children with cochlear implants. *Journal of Phonetics* 71 (2018) 98–112.

MacDonald, E. N., Johnson, E. K., Forsythe, J., Plante, P., & Munhall, K. G. (2012). Children's development of self-regulation in speech production. *Current Biology*, 22(2), 113–117.

Ménard, L., Perrier, P., Aubin, J., Savariaux, C., & Thibeault, M. (2008). Compensation strategies for a lip-tube perturbation of French [u]: An acoustic and perceptual study of 4-year-old children. *The Journal of the Acoustical Society of America*, 124(2), 1192–1206.

Vihman, M. M. (2017). Learning words and learning sounds: Advances in language development. *British Journal of Psychology*, 108(1), 1–27.