

Vowel duration before a geminate in children and adults: is lexicality important?

Lucia Colombo* and Joanne Arciuli^o

*Università degli Studi di Padova, Italy

^oFlinders University, Adelaide, Australia

ABSTRACT

The phonetic realization of Italian vowels includes shorter vowel durations before geminate consonants compared to singleton consonants but only in stressed syllables (/p'ettine/, comb). Previous research using nonword repetition showed that Italian children reach adult-like mastery of the vowels production in such contexts after 4 years of age [7]. We investigated whether this developmental trajectory holds when children produce familiar words. Italian children (3-6 years old) and adults performed a picture naming task, eliciting productions of familiar words. Stimuli were two and three-syllable words labelling familiar objects or people, with medial singleton or geminate consonants, and with stress on the initial or penultimate syllable. Acoustic analyses showed overall longer durations of vowels preceding the medial consonant (V1) for young children compared to adults. In contrast to the previous study of nonword production, even the youngest children were able to produce the appropriate adult-like variation in V1 duration in singleton/geminate contexts.

Keywords: Vowel production, Geminate, Lexical stress, Lexicality, Language acquisition.

1. INTRODUCTION

Geminate consonants occur frequently in some languages (e.g., Arabic, Finnish, Italian). Phonetically, they are realized as long consonants. Acoustically, a geminate consonant has a longer duration by comparison with the singleton consonant (e.g., the minimal pair in Italian *note*, notes, vs *notte*, night). In Italian, and some other languages, the realisation of vowels that surround medial consonants differs depending on whether the medial consonant is singleton or geminate [9;12]. In particular, in Italian, the duration of the vowel preceding a geminate is shorter compared to when it precedes a singleton. This difference in vowel duration, that is dependent on the length of the following consonant, is useful for the listener, to better discriminate the geminate/singleton contrast [9; 11]. However, this shortening of the preceding vowel only occurs for tonic vowels, that is, where the vowel preceding the medial

geminate/singleton consonant carries stress [4; 7]). When this vowel does not carry stress, its duration is not shortened. In the present study we investigated the impact of these contextual factors of type of consonant and stress on vowel duration in the production of real words by children and adults.

1.1. Stress in Italian

In Italian vowel duration depends on different contextual factors, one of which is stress. Most Italian words are stressed on the penultimate syllable (~75% of words) (e.g., /bam'bino/, child; /'kane/, dog), while a smaller number of words have stress on the antepenultimate syllable, (~18% of words (e.g., /'tavolo/, table) [6; 16]. Stress on the final syllable is much less frequent (about 2-3%). The main acoustic correlate of stress in Italian is the increased duration of the vowel, compared to a non-stressed syllable, in particular in open non-final syllables [3; 8]. However, the phonetic realization of the vowel preceding the medial consonant is shorter when it precedes a geminate compared to a singleton (e.g., V1 is shorter in CV1CCV2 than in CV1CV2, where the first syllable is stressed: compare /'notte/, night vs /'note/, notes). This poses no problem for the adult native speaker, but for young children there is a developmental trajectory towards adult-like mastery of this aspect of speech production. That is, children must acquire the distinction between geminate and singleton consonants, as well as variable vowel duration related to the nature of the following consonant, as well as stressed vs. unstressed conditions. Italian children show adult-like mastery in terms of variable vowel duration in stressed and unstressed syllables, even at quite a young age [1; 13]. Recent evidence, however, shows that this is not so in productions that involve the medial geminate/singleton distinction. Note that we distinguish between intelligibility and adult-like mastery. Thus, children's productions may be intelligible but not yet adult-like as revealed by fine-grained acoustic analyses.

1.2. Children's acquisition of stress and geminates

Three- and four-years-old children may not show adult-like mastery of the geminate/singleton contrast [7], which is associated to the duration of both the

consonant, and of the preceding vowel. This has been interpreted as meaning that in children aged 3–4 years the contrast between geminate and singletons, and the ability to phonetically realize this distinction, has not been fully acquired (i.e., is not adult-like).

These results have been shown in a task that involved measuring vowel duration produced via repetition of nonwords which exhibited the geminate/singleton contrast in medial position [7]). In the present study we report the results of an experiment where children were required to name pictures of real objects (i.e., real word production). The aim was to investigate if adult-like mastery of the production of vowels preceding the singleton/geminate contrast depends on the lexical status of the stimuli (i.e., when familiar words are produced as opposed to nonwords). Indeed, the phonological representation of known words may be stronger and more stable [17] such that the phonetic realization of vowels preceding medial geminate/singleton consonants might be easier to acquire for young children by comparison with their productions of vowels in nonwords.

To test this hypothesis, we selected a number of pictures whose names were: three-syllabic words with penultimate syllable stress and medial singleton, (/ma'tita/, pencil) or medial geminate consonants (/tak'kino/, turkey); disyllabic and three-syllabic words with initial syllable stress and medial singleton, ('note/, notes; /sɛdano/, celery) or medial geminate consonants ('gatti/, cats; /pattino/, skate shoe). We expected the duration of the vowel preceding a medial singleton consonant to be longer than the vowel preceding a medial geminate consonant in words where this vowel carries stress (i.e., in words like /'note/, notes, and /sɛdano/, celery, compared to words like /gatti/, cats, and /pattino/, iceskate). This prediction should hold for adults and older children, as it did for the repetition of nonwords in Colombo et al.'s [7] study. The crucial question was whether young children (3 and 4 years of age) would be able to realize vowels differently depending on these contextual factors as adults do: when repeating nonwords they did not show adult-like mastery but we might see a different pattern of results when children are naming familiar words.

2. METHOD

2.1. Participants

Overall, 77 children from a pre-school in a small town in north-eastern Italy were recruited for the purposes of the research. Due to technical issues, only the recording from 74 participants were analyzed (42 female and 32 male). The age of the

participants ranged from 3 to 6 years: 17 children aged 3 years (mean = 3.62, SD = 0.22), 24 children aged 4 years (mean = 4.54, SD = 0.23), 21 children aged 5 years (mean = 5.58, SD = 0.28) and 12 children aged 6 years (mean = 6.22, SD = 0.18). The children were assigned both the word stimuli of the present study and a task of nonword repetition [7]. The recordings were undertaken individually and in a quiet room of the pre-school building. Only children whose parents signed an agreement form participated to the research. Additionally, 28 adults from north-east Italy were recruited, ranging from 19 to 62 years of age (mean = 33.89, SD = 13.94, 16 females and 12 males). Like the children, the adults were interviewed and recorded in an isolated and quiet room.

2.2. Materials

Several constraints were considered for the word-picture pairs selection. First, the objects corresponding to the words had to be easy to name by young children as well as adults. The tonic syllable of the words must be the first or second syllable. The CV structure of the words had to be as simple and common as possible, to avoid confounds due to structure complexity (in Italian the most common structure is CVCVCV).

Given that the majority of Italian words have penultimate stress, it was quite difficult to find three-syllabic words with initial stress associated with easily visualised familiar objects. Forty-three 2- and 3-syllable words with penultimate or initial stress and medial singleton or medial geminate consonants were initially selected. The CV structure was mostly CVCV- CVCCV for disyllables and CVCVCV for 3 syllables. Words corresponded to common objects, familiar to the younger children. They were photographs of real objects/people, found on the web. Out of the 43 words, 20 were three-syllabic with penultimate stress, 13 of which with a geminate consonant, and 7 with a medial singleton consonant. Five three-syllabic words had initial syllable stress, 3 of which had a medial geminate, and 2 a singleton. Finally, there were 18 disyllabic words, 17 of which with initial stress and a medial geminate. The pictures corresponding to these words were selected and presented in a pilot test to 11 adult participants for picture agreement verification. Pictures which did not elicit the expected name for the picture by at least 50% of the adult participants were discarded.

On the basis of the pilot, some words were excluded from the statistical analysis (mostly because they contained a diphthong after the medial consonant or because of low naming agreement.) After the exclusion of words that did not correspond to the

above-mentioned criteria, the following stimuli remained for the statistical analyses: 13 words with the stressed vowel before the medial consonant (V1) (6 with singleton, 7 with geminate consonant) and 13 words with an unstressed vowel (10 with geminate, 3 with singleton consonants). The corresponding pictures were presented in the naming task in order to elicit speech which could be analysed acoustically.

2.3. Procedure

The naming task was undertaken on a personal computer which displayed each picture embedded as a slide in a PowerPoint presentation. Each test slide was followed by an empty slide. The slide sequence was controlled by the experimenter. Four practice pictures, not part of the experimental materials, preceded the test stimuli. Two different random sequences of the stimuli were assigned to two groups of participants. The participants' responses were recorded with a microphone headset (Reloop RUF 1NH/HS) connected to a computer and with the software Audacity. The original (.aup) format was converted to another format (.wav 32 bit) for the acoustic analyses.

The software PRAAT, version 6.0.28, [5] was used to undertake acoustic measurements of vowels. Waveforms and wide-band spectrograms with a 300-Hz bandwidth were generated for each sound file. Vowel segmentation was made considering concurrent information from amplitude traces, intensity curves and F0 contour. Both the vowel preceding (V1) and the vowel following the medial consonant (V2) were analyzed. The onset of V1 was determined by an increase in amplitude and appearance of the formant structure. The offset of V1 was determined by a drop in amplitude and change in formant structure. The onset of V2 was determined by the release burst of the preceding stop consonant and appearance of formant structure, and its offset by an intensity variation in the waveform and spectrogram. All analyzed vowels were intermediate vowels that preceded a medial consonant (either singleton or geminate consonant).

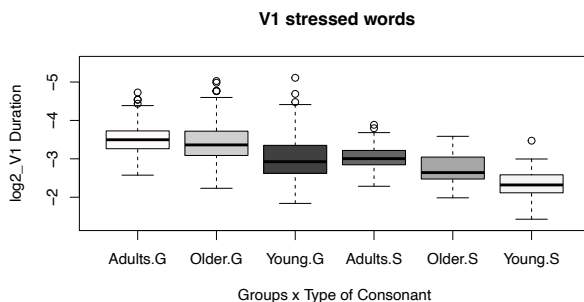


Figure 1: Log2 V1 duration in the stressed vowel V1 in the three groups for words stressed on V1.

3. RESULTS

Perceptual evaluations were made by the first author and the experimenter (both native speakers of Italian) in order to determine whether each production was correct and analyzable. Productions were considered correct if all segments were pronounced correctly, and with correct stress and distinctive single/geminate pronunciations, and if the utterance corresponded to the expected name of the picture. For two of the picture stimuli two names were possible and both were analyzed, when appropriate, one of which was a diminutive (/gatti/, /ga'ttini/; cat, small cats; /tattsa/, /tatts'ina/; cup, small cup). Both words of each pair had a medial geminate consonant, with V1 unstressed in the diminutive words.

Acoustic analyses were only undertaken on correct productions. For children, 1518 out of 1920 productions (79.06%) and for adults 668 out of 669 (99.85%) were available for acoustic and statistical analyses. The data from the children aged 3-6 were sorted into two age groups: 3-4 and 5-6, so as to have a simpler statistical design.

The analyses of the dependent variable of vowel duration were conducted using linear mixed-effects modelling in R version 3.5.2 GUI [14], treating participants and items as random effects. Models with random slopes in addition to random intercepts by participants showed convergence problems, so only random effects for intercepts were included. On the basis of the reviewed literature, we expected an effect of the type of consonant only on stressed V1 (not unstressed V1), thus the data for the two types of stimuli were analyzed separately. Fixed factors were group with three levels (3-4, 5-6, and adults, between-subjects and within-items), and type of consonant (singleton/ geminate, within-subjects and between-items). The models were fit by maximum likelihood with the Laplace approximation technique. The lme4 package, version 1.1-21 [2] was used to run the linear mixed-effects models. As no effect was found in the analyses of V2 durations, and none was predicted, only the results of the analyses on V1 durations will be reported. The analyses were carried out on log2 transformed data, in order to ensure a normal distribution.

3.1. V1 in stressed syllables

We used a statistical model with a two-way interaction: $\text{lmer}(\log_2_V1_Duration \sim \text{Group} * \text{Type of consonant} + (1|\text{Participant}) + (1|\text{Word}))$, and it explained more variance compared to the model with only the additive effects, $\chi^2 = 9.93$, $p = .001$. The ANOVA showed significant effects of group, $\chi^2 = 228.01$, $p = 0$, type of consonant, $\chi^2 = 17.27$, $p = 0$,

and of the interaction, $\chi^2 = 9.98$, $p = .001$. The group effect was due to decreasing V1 durations in young children by comparison with adults (Figure 1). The effect relating to type of consonant was due to longer durations of V1 preceding a singleton than V1 preceding a geminate. Separate contrasts between the means for vowel duration in singleton/geminate contexts for each group showed that the difference in vowel duration for singleton vs. geminate contrasts was significant for each group, but was numerically larger for younger children due to the longer overall V1 duration: $t(22.8) = 5.26$, $p < .0001$; $t(19.3) = 5.43$, $p < .0001$; $t(118.3) = 3.90$, $p = .001$, respectively for young children, older children and adults. Analyses conducted on just three-syllabic words gave a similar pattern of results, with longer durations of V1 preceding a singleton than preceding a geminate for all the groups.

3.2. V1 in unstressed syllables

The means for the unstressed vowel are reported in Figure 2. In contrast to the former analysis, the statistical model including the interaction did not significantly differ from the model with only the two fixed factors. The ANOVA was carried out on the model with two fixed factors, group and type of consonant, and the random factors, participants and words. Only the group effect was significant, with decreasing V1 durations from younger children to adults: $\chi^2 = 289.39$, $p = 0$. Comparisons of the vowel durations of all the groups were significantly different: $t(874) = 16.37$; $t(629) = 8.88$; and $t(282) = 5.24$, all p 's $< .0001$, respectively for the young and the older children compared to adults, and for the two children's groups.

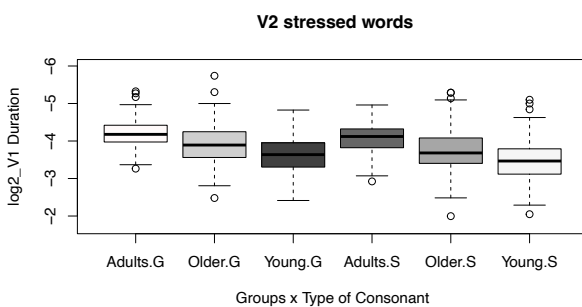


Figure 2: V1 duration in msec in the unstressed vowel V1 in the three groups for words with stress on V2.

4. DISCUSSION

The results of the experiment using speech production elicited via picture naming confirmed that, in Italian, vowel duration in stressed vowels before a geminate is shorter than before a singleton in adults [1; 7]. In

the groups of children we tested, the same pattern was found, even though speech rate is slower in children, and therefore vowel duration was overall longer than in adults. In contrast to previous research that elicited speech via nonword repetition, even young children are able to realize vowels differently depending on the contextual factors of a following singleton/geminate, and on whether the vowel occurs in a stressed syllable or not. Note that the participants in the present study were the same as in the previously published study with nonwords [7]. As has been discussed in the latter paper, this contrast between production of words and nonwords might depend on a variety of factors. One is the familiarity associated with real words. For example, children aged 26 months are able to repeat familiar words better than nonwords even when the words include unfamiliar consonants [10; 15]. This in turn might depend on different factors. First, children may learn words, and thus produce them, as units [17]. Moreover, learned words are retrieved from memory, where their phonological representation is also connected to semantics, which may act as support to the phonological representation. Indeed, the interconnections between the word phonology and its meaning may reinforce the phonological representation held in the memory buffer.

In contrast, nonwords are sequences of segments never heard before. In order to be repeated, they must be held in a memory buffer, while the production processes, required to articulate them aloud, are activated. Further, nonwords are presumably represented as sequences of phonemes or groups of phonemes in a specific order. The motor program to articulate them is based on this memorized sequence from short-term memory.

In conclusion, in Italian, realization of vowels duration is affected by a number of contextual factors including length of the following consonant and whether the vowel occurs in a stressed syllable or not. The present study confirms that lexicality is another crucial factor in the acquisition of this contrast. However, the present results do not allow us to determine at what level of representation or implementation the failure to produce the singleton/geminate contrast in V1 duration is located (whether it is perceptual, at the level of phonological or phonetic representation, or even at the motor production level).

In further studies it will be important to investigate whether this word/nonword difference only depends on familiarity of the word to pronounce, or also from other factors, like its meaning.

5. REFERENCES

- [1] Arciuli, J., Colombo, L. 2016. An acoustic investigation of the developmental trajectory of lexical stress contrastivity in Italian. *Speech communication*, 80, 22- 33.
- [2] Bates, D., Mächler, M., Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.
- [3] Bertinetto, P.M. 1980. The perception of stress by Italian speakers. *Journal of Phonetics*, 8(4), 385-395.
- [4] Bertinetto, P.M. & E. Vivalda. 1978. Recherches sur la perception des oppositions de quantité en italien. *Journal of Italian Linguistics* 3, 97-116.
- [5] Boersma, P., & Weenink, D. (2017). Praat v. 6.0.28 [Computer software]. Amsterdam: Institute of Phonetic Sciences.
- [6] Colombo, L. 1992. Lexical stress effect and its interaction with frequency in word pronunciation. *Journal of Experimental Psychology*, 18, 987-1003.
- [7] Colombo, L., Infanti, M., Arciuli, J. 2022. How is vowel production in Italian affected by geminate consonants and stress patterns? *J Child Lang*. 2022 Nov 3:1-19. doi: 10.1017/S0305000922000526.
- [8] D’Imperio, M.R., & Rosenthal, S.S. 1999. Phonetics and phonology of main stress in Italian. *Phonology*, 16, 1-28.
- [9] Esposito, A., & Di Benedetto, M. G. (1999). Acoustical and perceptual study of gemination in Italian stops. *Journal of the Acoustical Society of America*, 106(4, Pt 1), 2051–2062. <https://doi.org/10.1121/1.428056>
- [10] Keren-Portnoy, T., Vihman, M. M., DePaolis, R. A., Whitaker, C. J., Williams, N. M. 2010. The role of vocal practice in constructing phonological working memory. *J Speech Lang Hear Res*. 53(5), 1280–1293. [https://doi.org/10.1044/1092-4388\(2009/09-0003\)](https://doi.org/10.1044/1092-4388(2009/09-0003))
- [11] Krull D., Traunmüller, H. & Bertinetto, P.M. (2006). Local speaking rate and perceived quantity: an experiment with Italian listeners. *Working Papers* 52 (Lund University, Dept. of Linguistics), 81-84.
- [12] Maddieson, I. 1985. Phonetic cues to syllabification. In: *Phonetic linguistics: Essays in honor of Peter Ladefoged*, Academic Press, 203-221.
- [13] Olivucci, F., Pasqualetto, F., Vayra, M., & Zmarich, C. (2016). Lo sviluppo dell’accento lessicale nel bambino in età prescolare: una prospettiva fonetico- acustica. In R. Savy & I. Alfano (cur.), *La fonetica nell’apprendimento delle lingue* (pp. 209-228). Milano: Officinaventuno.
- [14] R Core Team (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org>
- [15] Roy P, Chiat S. A prosodically controlled word and nonword repetition task for 2- to 4-year-olds: evidence from typically developing children. *J Speech Lang Hear Res*. 2004 Feb;47(1):223-34. doi: 10.1044/1092-4388(2004/019). PMID: 15072541.
- [16] Spinelli, G., Sulpizio, S., & Burani, C. (2017). Q2Stress: A database for multiple cues to stress assignment in Italian. *Behavior Research Methods*, 49(6), 2113–2126. <https://doi.org/10.3758/s13428-016-0845-7>
- [17] Vihman, M., & Keren-Portnoy, T. (Eds.). (2013). *The Emergence of Phonology: Whole-word Approaches and Cross-linguistic Evidence*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511980503