DOI: 10.17469/O2103AISV000017

CLAUDIO ZMARICH, ELISA RACCANELLI, GIULIA RANCAN, CARLA LEONARDI

Instability of speech production as a marker of Childhood Apraxia of Speech (CAS): Segmental and acoustic evidence

According to ASHA (2007), CAS is "a neurological childhood disorder in which the precision and consistency of movements underlying speech are impaired in the absence of neuromuscular deficits". The present study aims to verify whether the extreme variability (at segmental as well as acoustic level) in the production of the very same linguistic unit (word or syllable) is a marker of CAS. We engaged three CAS subjects in multiple productions of the same linguistic units using the TFPI, a new phonetic test not yet published. Then we analyzed the recorded signals by means of different metrics, chosen among the most sensitive ones to track speech variability, in order to assess the consistency and stability of CAS subjects production. This was compared to that of the control groups, i.e. lexical age peers (calculated with the Italian version of MacArthur Communicative Development Inventories or CDI), chronological age peers and adults. The results suggest that CAS speech is characterized by phonological inconsistency, i.e., multiple productions of the same target very often do not share the same phonological form, especially with the longest words. Moreover, the analysis of intra-syllabic CV anticipatory coarticulation suggests that CAS children have an immature speech motor control, similar to that of 2- or 3-years-old children.

Key words: Childhood Apraxia of Speech (CAS), Coarticulation.

1. *Introduction*

Childhood Apraxia of Speech (CAS) is a neurogenic Speech Sound Disorder (SSD; DSM-V, APA 2013) whose etiology and neurobiological correlates are still unclear. According to ASHA (2007), CAS is "a neurological childhood disorder in which the precision and consistency of movements underlying speech are impaired in the absence of neuromuscular deficits", that is, a speech motor disorder whose core deficit involves the planning and/or programming of the spatiotemporal parameters of movement sequences (Terband, Maassen, 2010). From a segmental point of view, CAS subjects often produce a reduced phonetic inventory than expected for their chronological age, and the production is often characterized by inaccuracy in the realization of speech sequences, resulting in phonemes omissions, additions, substitutions and distortions that frequently make the speech unintelligible. Adding to this, segmental errors increase as the words get longer, and/or the phonemes are part of an accented syllable (Chilosi, Lorenzini, Cerri & Cipriani, 2014). The speed of articulation of CAS subjects, measured in number of syllables per second, is significantly slower than their peers (Chilosi, Lorenzini, Fiori, Graziosi, Rossi, Pasquariello, Cipriani & Cioni, 2015). CAS subjects may also present several markers of morphological and syntactic disorders. Syntactic structures suffer from the limitations of the vocabulary and CAS children tend to produce sentences of less than average length and with a simpler grammatical structure compared to their peers (Chilosi *et al.*, 2015). It has been hypothesized that the small vocabularies of children with CAS have an upstream impact on the cognitive-linguistic aspects of phonology (Velleman, 2011).

Considering the suspect of an underlying deficit in planning or programming speech movements, it is not surprising that coarticulation in CAS subjects has been thoroughly investigated in recent years, through both acoustic and kinematic analysis. Acoustic analysis of CAS speech production was performed by means of two different methodologies, mainly by L. Nijland and colleagues on the one hand and H. Sussman and colleagues on the other.

In a first experiment, Nijland, Maassen, van der Meulen, Gabreëls, Kraaimaat & Schreuder (2002) investigated the magnitude and variability of the anticipatory coarticulation exerted by the last vowel on the first vowel and on the consonant of the nonsense disyllables [əCV], where V was /a i u/ and C was /s x b d/. Results showed that CAS children were significantly more variable in the F2 trajectories than the age matched control children and the adults, and they exhibited less anticipatory coarticulation than their controls. In a second experiment (Nijland, Maassen, van der Meulen, Gabreëls, Kraaimaat & Schreuder, 2003), CAS children and control children produced high- and low-frequency of occurrence syllable utterances, in which the syllable structures were systematically manipulated, letting unchanged the phonemes sequence (-V1s#xV2 vs -V1#sxV2). Anticipatory coarticulation, using second formant trajectories, and durational structure were analyzed. This time the results showed stronger coarticulation in CAS children when compared to controls. Furthermore, at the prosodic level, CAS children, differently from controls, did not show any metrical contrast effects, i.e. they do not change the durational structure of V1 and /s/ according to changes in stress level and/or syllabic structure.

The study by Sussman, Marquardt & Doyle (2000) is exemplificative of the second methodology. The authors made use of the Locus Equation analysis of coarticulation, of which the first author is the main proponent (for a description of the method, see the procedural section of the present paper). The authors measured the degree of anticipatory coarticulation in CV monosyllabic words produced on imitation by a small group of CAS children and a group of age matched controls. Once compared to controls, CAS *k* values (the slopes of the Locus Equations indexing the degree of coarticulation) revealed a reduced distinctiveness in intra-syllabic coarticulatory extent across stop place categories (labial, dental, velars), and lower R2 and larger SE values, indicative of greater variability.

The purpose of the present study is to verify whether the extreme variability at the segmental (defined as phonological inconsistency) as well as at the acoustic level (measured as degree of anticipatory coarticulation) in multiple productions of the very same linguistic unit, respectively word or syllable, could be a marker of CAS, and, as to coarticulation, whether the CV syllables produced by CAS children are more or less coarticulated than those produced by their controls[1].

2. *Method*

2.1 Subjects

Three subjects were reported to be affected by CAS by clinicians (one of them being the last author of the present paper) (see Table 1).

The first one (GE) presented a lexical comprehension level slightly below the chronological age (PPVT, Stella, Pizzioli & Tressoldi, 2000), while the verbal production showed a serious deficit. Morphosyntax was seriously under the standard scores both for comprehension (TCGB, Chilosi, Cipriani, 1995; PVCL, Rustioni Metz Lancaster, 2007) and production (PVB, Caselli, Pasqualetti & Stefanini, 2007). The test scores for nonverbal oral and face movements showed that also this area was seriously affected, with more difficulty on performing praxias[2] on imitation than on request (Bearzotti e Fabbro, 2003). The Phonological Working Memory (PWM) could not go over the two elements (VAUMeLF test, Bertelli, Bilancia, 2008).

Table 1 - *Clinical characteristics of CAS subjects*

	GE	*RA*	*DO*
AGE (years.months)-SEX	11.0 - F	10.4 - F	8.3 - M
Lexical Production (CDI, Caselli *et al.*, 2007)	623 words (lexical age of slightly more than 36 months)	445 words (lexical age of 31 months)	617 words (lexical age of slightly more than 36 months)
Lexical Comprehension (PPVT, Stella *et al.*, 2000)	Slightly under chronological age	Slightly under chronological age	Under chronological age
Morphosyntax Production (CDI, Caselli *et al.*, 2007)	Seriously under chronological age	Seriously under chronological age	Seriously under chronological age

[1] Inconsistency is a clinical marker of CAS if it includes some important signs: i) inconsistent errors with groping (a silent research of the articulatory locus made by tongue, lips and jaw) for consonants and vowels during the production of syllables or words; ii) unpredictable articulatory production (i.e., consonant and vowels are performed in various ways both in different words and in the same word); iii) erratic errors, atypical and not recurrent phonological processes; iiii) efforts characterized by groping. Sometimes CAS speakers make repetitions or breaks that can be wrongly interpreted as stuttering. Interestingly, inconsistency has been observed in many languages such as Cantonese, Turkish, German, Portuguese.

[2] Praxia (from ancient Greek *prásso* «act») is an intentional action that requires the ability to plan, program and execute a sequence of movements to reach a purpose or an objective.

Morphosyntax Comprehension (Rustioni *et al.*, 2007)	Seriously under chronological age	Seriously under chronological age	Seriously under chronological age (TROG-2, Bishop 2009)
Nonverbal Oral & Speech Motor Control (Bearzotti, Fabbro, 2003)	Seriously affected (more on imitation)	Seriously affected (more on imitation)	Seriously affected (more on request)
Phonological Working Memory (VAUMeLF, Bertelli, Bilancia, 2008)	No more than two elements	Just 1 element	No more than two elements

The second subject (RA) presented a lexical comprehension level, according to PPVT, below the standard scores, while the lexical production showed a serious deficit. Morphosyntax was seriously below the standard scores both for comprehension (TCGB, PVCL) and production (PVB). The test scores for nonverbal oral and face movements (Bearzotti, Fabbro, 2003) revealed these skills to be seriously affected, with more difficulty when performing praxias on imitation than on request. The PWM scored 1 (very low). Articulatory diadochokinesis (Williams, Stackhouse, 1998) was also severely impaired.

The third subject (DO) presented lexical and morphosyntax comprehension levels under the standard scores (PPVT; TROG-2, Bishop 2009). The test scores for nonverbal oral and face movements (Bearzotti, Fabbro, 2003) evidenced a difficulty in performing praxias, more on request than on imitation. Articulatory diadochokinesis (Williams, Stackhouse, 1998) was also severely affected. The PWM was limited to two elements (VAUMeLF).

Control subjects were recruited according to three different criteria. A first group of children, raised in a monolingual context and developing in a typical way according to their parents, were recruited because they scored the same number of words at the Italian version of the MacArthur Communicative Development Inventories or CDI (Caselli *et al.*, 2007) as the CAS subjects (lexical age peers: ME, male, 3 years, 6 months, 633 words attested on CDI; NL, female, 2 years, 7 months, 445 words attested on CDI; DBN, male, 3 years, 2 months, 610 words attested on CDI).

Two other groups were considered for comparison: 3 typically developing children (according to parents' statements), which have the same chronological ages as the CAS subjects (chronological age peers: MA, female, 11 years, 0 months; FG, female, 10 years, 4 months; FS, male, 8 years, 3 months), and 3 adult subjects, all females because their F0 is more similar to children's F0 (HC, 28 years; MF, 23 years; PM 21 years), who declared not to have ever suffered for cognitive, speech or motor impairment and to speak Italian as their native language.

2.2 Procedure

The aim of the present study is to verify whether the extreme variability in the production of the very same linguistic unit (word or syllable) could be a marker of CAS. Consequently, we devised to engage the subjects in multiple productions of the same word, and then we analyzed the recorded signals by means of different metrics (both segmental and acoustic), chosen among the most sensitive in tracking variability, in order to compare the consistency/stability of the CAS subjects to that of their control subjects.

CAS subjects were administered the Italian version of CDI (Caselli *et al.*, 2007) which consists of a 680-words vocabulary production checklist, to be compiled by parents, which is designed for use with children in the 1;4-2;6-years range. It allowed us to calculate the lexical age, based on total words produced according to their parents (the same test, as already said, was administered to the younger control group). According to CDI, CAS subjects resulted to have a relatively restricted lexicon (GE: 623 words; RA: 464 words; DO: 617 words).

CAS subjects and control subjects filled (or the parents filled for them, according to the ages of the subjects) a custom-made, socio-linguistic questionnaire, which reported basic information about psycho-physiological and linguistic development, and they performed the Test Fonetico per la Prima Infanzia (TFPI), a new phonetic test not yet published (see Zmarich, Fava, Del Monego & Bonifacio, 2012). The TFPI contains a naming task divided into two subtests based on the age of the subjects (18-23 months; 24-47 months). Its construction took into account a Phonetic criterion, i.e. the consonant phonemes of Italian must be established in at least two different words for each position of the tested words, a Semantic/Frequency criterion, i.e. the words have to be high frequency concrete nouns, and a criterion of Gradualness in phonetic complexity, i.e. words must evidence an increase in complexity from the first to the second age group, by number and types of syllables. The TFPI form appropriate for subjects aged 25-47 months was chosen, which includes 78 figures displayed on a PC screen to be named by the subjects. Considered as a group, these words include all the 5 vowels of the Italian unstressed vowel system [i e a o u], and each vowel occur more than twice. This particular was critical in order to ensure variability in the vocalic segments, which is a pre-condition for the acoustic analysis of coarticulation by means of the Locus Equations method[3]. Crucially, in order to create opportunity for speech variability manifestation, the subjects had to produce the TFPI items three times in the same day (each time the order of items presentation was randomized) and the lexical productions were recorded at 44 kHz and 16 bit by means of an Edirol R-09 in a silent room.

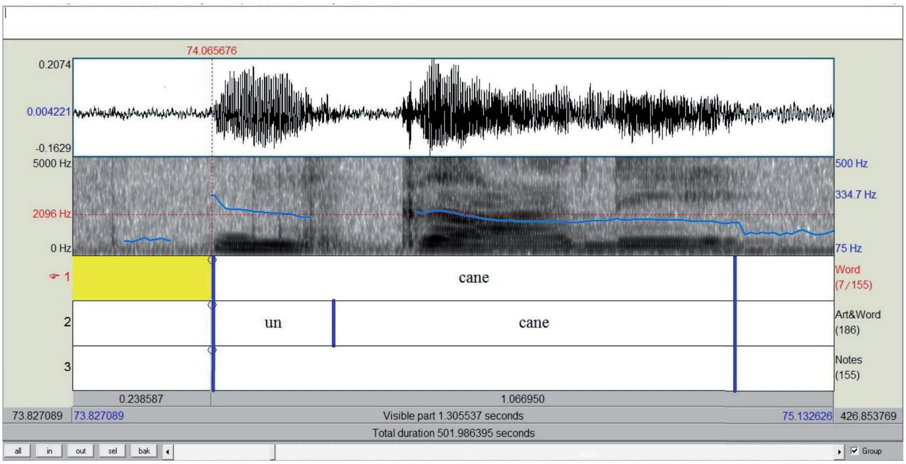
Phonetic transcription (in IPA) was always performed by the second or the third author working in an exclusive way on a single subject (i.e., each subject has been transcribed by the same transcriber). All the transcriptions were then checked

[3] For this reason, we made sure that all the subjects produced the syllables beginning with plosives with a statistically sufficient number of occurrences for at least the cardinal vowels [i a u].

by the first author, which took also the final decision as to the few disagreements. During the transcription process, the PRAAT software (http://www.fon.hum.uva.nl/praat/) was used as a practical help and also to segment and label the words produced by subjects. Only the fluent and intelligible productions of correct or incorrect words of TFPI were considered for the analysis.

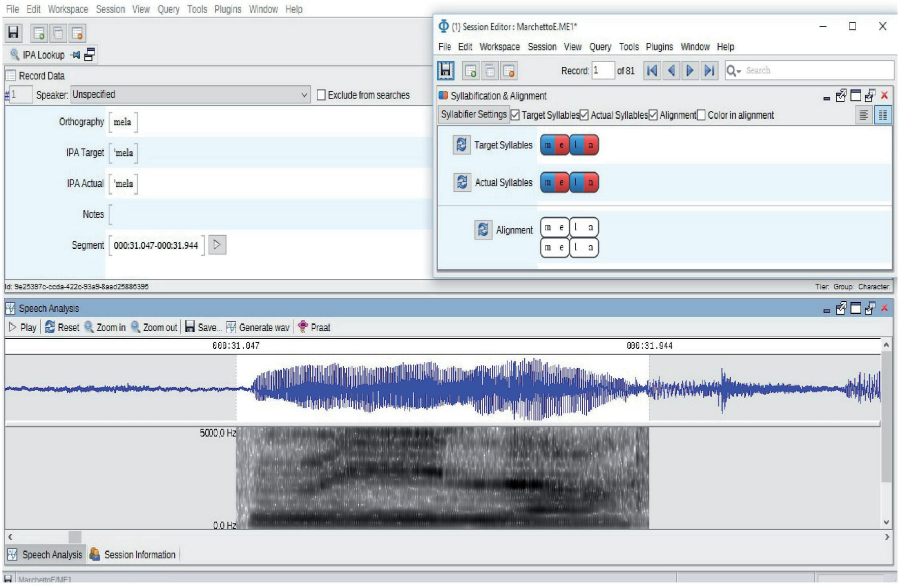
After that, words in the Word tier of PRAAT (see Figure 1) were exported as records to a PHON database (https://www.phon.ca/phontrac; Rose, Stoel-Gammon, 2015), in order to accomplish the phonological analysis made possible by the syllabification, segmental alignment, and matching between each lexical target (IPA Target) and the child's effective realization (IPA Actual, see Figure 2).

Figure 1 - *Praat display reporting the waveform of the lexical production segmented and annotated according to three tiers*



To this end, we exploited the interfacing between PRAAT and PHON to convert the content and the timing information, represented in the PRAAT Word tier by the orthographic word and by the initial and final border, respectively, into a record of Phon, by means of a custom script written by Vincenzo Galatà (CNR-ISTC). A PHON record is the basic element of a PHON database, and in this case is represented on the screen by an orthographic transcription of the target word (the word the child is attempting to), an IPA transcription of the target word, and an IPA transcription of the child attempt. Close to this section there is the acoustic signal of the word accompanied by the article, if present (corresponding to the Art/Word tier on PRAAT). In this way, a PHON database may contain a number of records corresponding to all the words produced by every single child. They appear already transcribed in IPA on the IPA Target tier thanks to an Italian phonetic dictionary incorporated into PHON. As a final step, the user of PHON will fill with IPA symbols the tier devoted to the effective child's production (IPA Actual). At this point data are ready to be analyzed through comparisons between target (adult/model) and actual (child/produced) form, and an algorithm that performs bestguess segmental alignments between corresponding target and actual IPA will help to classify the children's errors in one of the several types of phonological processes, or to produce a phonetic inventory, or the percentage calculation of consonants correctness, etc.

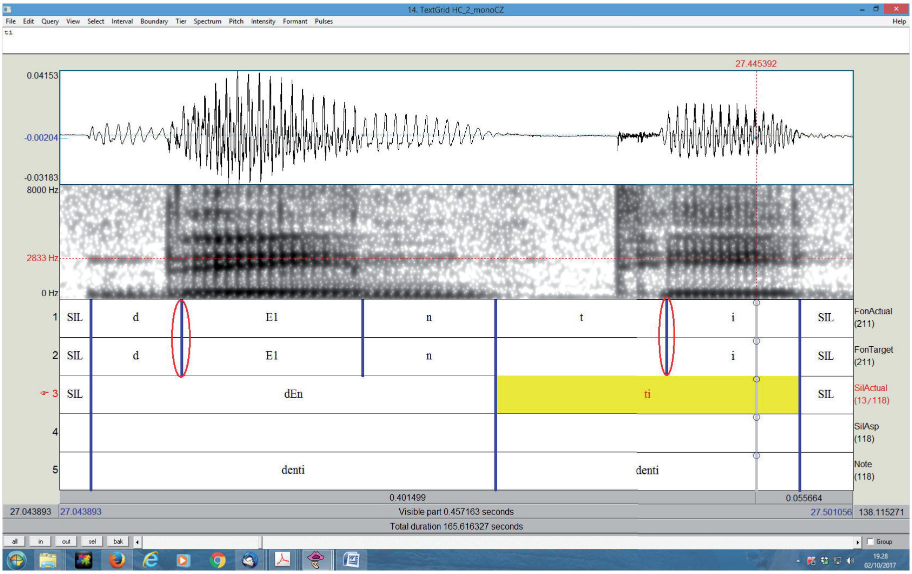
Figure 2 - *PHON display reporting the waveform of the lexical production (i.e. the record, at the bottom), three tiers containing the transcription of the lexical target (IPA Target) and the child's effective production (IPA Actual)*



Successively, any clear and intelligible C1V and C1VC syllables (C1 = plosive; V = any vowel) were looked for and further segmented and labeled at phone and syllable levels and processed by a PRAAT script (by V. Galatà) to obtain F2 values (Hz) on the first available oscillatory cycle after closure release (see "denti" in Figure 3) and at mid-vowel. A SilAsp tier was also included in order to report any aspirated stop release (more than 50 ms of aspiration after the release), which were later excluded from the coarticulation analysis for possible flaws. A second script calculated VOT, but we do not report the results of the VOT analysis here.

Since we would like to assess the degree of Consonant-Vowel intrasyllabic, anticipatory coarticulation, we calculated it by means of the Locus Equations method (Krull, 1988; Sussman, Marquardt & Doyle, 2000; see Sussman *et al.*, 1999, for an application to speech development), separately for bilabial, dental and velar place of articulation of the consonants.

Figure 3 - *PRAAT display: on the top, the waveform of a lexical production ("denti"), on the bottom, five tiers reporting the segmentations and labeling at different levels (word, syllables, segments). The boundaries used to extract F2 values at the C-V border are circled by a red line. The "SilAsp" tier is used to mark aspirated syllables, i.e. syllables beginning with a voiceless aspirated plosives (VOT greater than 50 ms)*



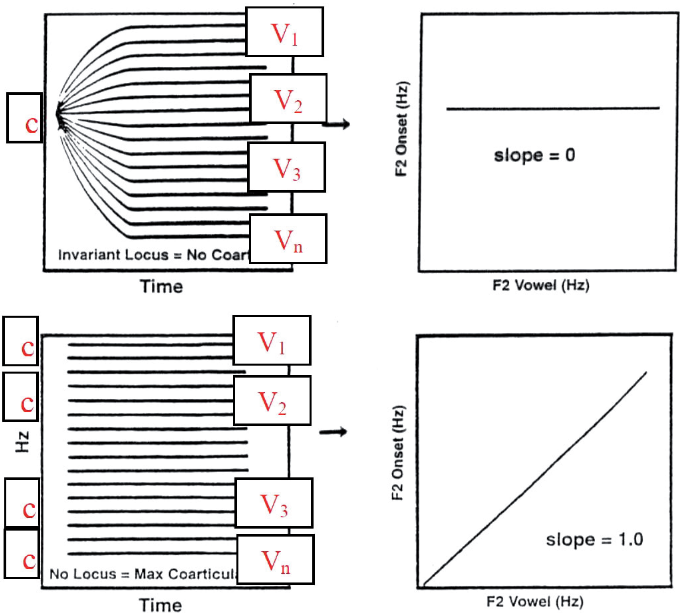
In the Locus Equations, the coefficient *k*, representing the slope of the regression line described by the equation: *F2 Onset = k * F2 Vowel + c*, indexes the degree of coarticulatory influence of V on C. The *k* coefficient and the *c* coefficient are calculated by means of a regression analysis of the dependent variable (the F2 consonant values) on the independent variable (the F2 vowel values). The value of *k* could vary between 0 (no coarticulation at all) and 1 (maximal coarticulation, Figure 4). It is interesting to note that F2 onset and F2 vowel, within a given place category, are consistently and robustly linearly correlated across diverse speakers (they are inherently normalized because they result from measures taken from the *same* syllable) and languages, and even under perturbation conditions as imposed by bite blocks. In addition, the particular linear function relating these two parameters is itself a function of place of articulation. Labials have been found to have the steepest regression functions, followed by velars, and then alveolars (for Italian, see Zmarich, Bortone, Vayra & Galatà, 2013).

3. *Results*

The first block of the results we present are drawn from Rancan (2015), which compared the segmental characteristics of the lexical productions of two CAS subjects (GE and RA) with those produced by the lexical age peers, on a number of phonetic-phonological measures (derived from the compilation of phonetic inventory and from the error analysis). We choose to match CAS subjects to lexical age peers because we felt that the alternative choice, involving a comparison with age matched normal subjects (from 8 to 11 years), would have been too "punitive" for CAS children.

The results from the acoustic analysis are drawn from Raccanelli (2016), which recruited another CAS subject (DO) and analyzed the lexical productions on a number of acoustic measures (VOT and degrees of intra-syllabic coarticulation). In the last work, CAS subjects were compared to lexical age peers, as before, but also to the chronological age peers and to the adults. In fact we felt that the previous cautions concerning the non-use of the age matched controls when using the phonetic test could be attenuated when a measure of speech motor functioning (coarticulation) was considered.

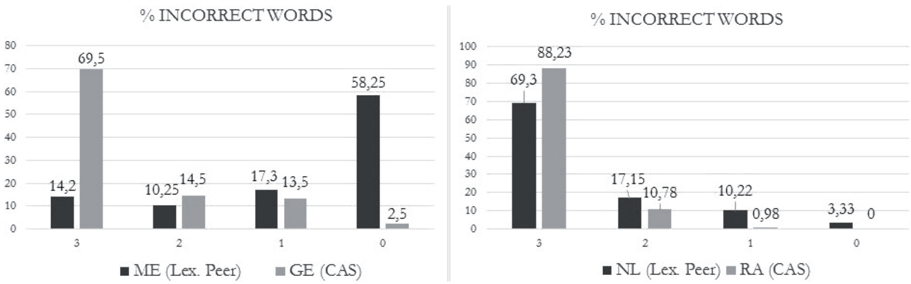
Figure 4 - *Hypothetical extremes of Locus Equation slopes (modified from Sussman et al., 1999). The top panel illustrates the F2 transition representing no coarticulation between the vowels and the consonants, and the Locus Equation slope of zero that would result from such a situation. The lower panel illustrates maximum coarticulation between vowels and consonants with no fixed consonantal locus and a resulting Locus Equation slope of 1.0*



3.1 Results of the segmental analysis

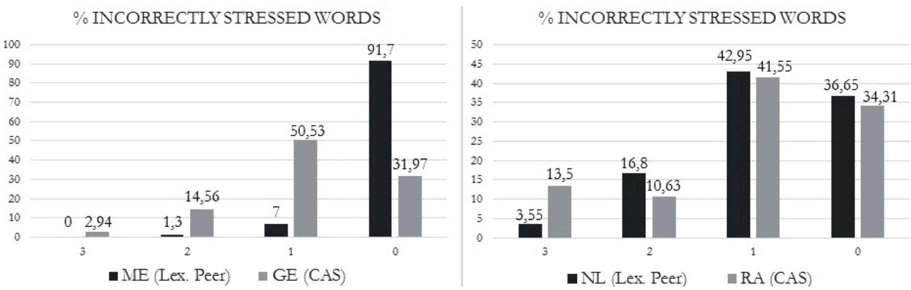
According to segmental analysis, the speech production of GE and RA sounds scarcely fluent and is affected by a wide range of error processes. In line with the literature findings, the subjects present a difficulty in combining phones in syllables and syllables in words. The percentage of errors increases with length and complexity of words. The subjects have an incomplete and atypical phonetic inventory, since they have not yet consistently acquired the consonants normally mastered since 24-32 months of age (Zmarich, Bonifacio, 2005). Speech is characterized by widespread abnormalities on different levels, in rhythm as well as in stress allocation. As far as consistency is concerned, we can select some analysis in order to demonstrate that multiple productions of the same speech targets are phonologically more inconsistent (from trial to trial) than those produced by the lexical age peers. Figure 5 shows the percentage distribution of incorrect words (due to segmental omissions, additions, substitutions and distortions), according to the number of words out of the totals, that were never mistaken (0), mistaken once out of three repetitions (1), mistaken twice (2), mistaken all the times (3). Although the difference is particularly evident for GE compared to the lexical peer ME, which produced incorrectly for three times the 69% of words, it is notable even for the other pair of subjects.

Figure 5 - *Percentage distribution of segmentally incorrect words according to the number of trials (3 = always incorrect; 0 = always correct)*



We know from the literature (i.e. Chilosi, Lorenzini, Cerri & Cipriani, 2014) that lexical stress is often displaced to another, normally unstressed, syllable. If we define a word as "incorrectly stressed" when we perceive that lexical stress is allocated on a syllable which is different from that one of the standard pronunciation, we found that one of the two CAS subjects (GE) stressed correctly all the repetitions of the lexical targets only the 31.97% out of all the attempts (Figure 6).

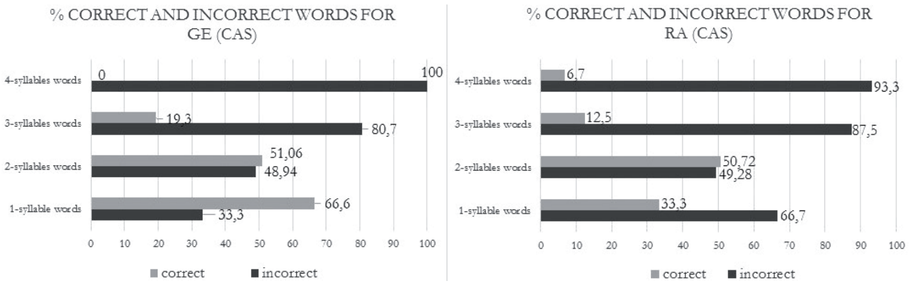
Figure 6 - *Percentage distribution of incorrectly stressed words according to the number of trials (3 = always incorrect; 0 = always correct)*



This low value contrasts with the 91% of total words always correctly stressed by the lexical peer (ME).

From the same literature we also know that longer words present greater difficulties for CAS subjects. Figure 7 shows the percentages of segmentally correct and incorrect words out of the total number of words having a given length in syllables (from one to four syllables). In this case, the total is represented by all the repetitions of the target words having the same length in syllables.

Figure 7 - *Percentage distribution of segmental correctness of the words, according to the number of constituting syllables, calculated over all the three trials*



The two CAS subjects perform similarly (but, again, GE is worse), generally making more and more mistakes as the words get longer[4]. However, since 4-syllable words have more segments than monosyllabic words, they are also more likely to contain more errors. In order to ascertain whether longer words are effectively more subject to errors, we calculated the number of incorrect syllables out of the number of the syllables in the lexical targets, by excluding errors relative to vocalism (because more prone to biases in the phonetic transcription), but including errors in syllabification (i.e. degemination, causing a CVC.CV word structure to become CV.CV). The results definitively excluded this possibility: as an example, the CAS subject GE, who was the worse, attempted to produce the 23 three-syllabic disyllabic target words three times each (for a total of 207 target syllables), and the resulting rate of incorrect/target syllables was 0.194 (almost one syllable incorrect out of five), which contrasts with the rate of 0.312 (almost one syllable incorrect out of three), when he attempted to produce the 50 disyllabic target words, three times each (for a total of 300 target syllables).

[4] There seems to be an exception: RA has more correct disyllabic than monosyllabic words. We can justify this behaviour by the current knowledge about the acquisition of speaking ability in CAS children: words and sentences tend to be acquired as whole units and the child can pronounce a whole word or even a sentence (such as "What time is it?") by the recruitment of an automatic mechanism. In this condition the articulation shows high rate fluency. On the contrary, when the child is trying to articulate intentionally rather than automatically, she/he may not be able to do it, even in the case of short syllables that frequently are codified as non-words. This situation influences the production by slowing down the fluency and/or leading to pronunciation errors. This phenomenon is called the automatic-voluntary dissociation. It is an erratic and highly variable behaviour, quite difficult to fully understand.

3.2 Results of the acoustic analysis

As already described in the procedure, the degree of anticipatory coarticulation of the vowel on the consonant, in all the CV and CVC syllables produced by the subjects in relation to the target words of the TFPI, has been calculated by means of Locus Equation method, separately for the bilabial, dental and velar place of articulation of the consonants. For this analysis, a third CAS subject was added, who was administered the phonetic test. This time, the control subjects were not limited to the lexical peers, but the chronological peers and the adults were also considered.

The experimental hypothesis consisted in verifying if, compared to controls: i) CAS subjects show similar degrees of coarticulation; ii) CAS subjects are able to differentiate the three places of articulation (i.e. are able to use a distinct degree of coarticulation for each place, as it is attested for Italian, see Petracco, Zmarich, 2006, and for other languages, see Sussman, McCaffrey & Matthews, 1991); iii) CAS subjects evidence more variability from trial to trial; iiii) CAS subjects evidence more variability in general, independently from trials.

The number of CV and CVC syllables of the intelligible lexical productions selected for each trial of each subject and categorized according to the articulatory place was around 20, ranging from a minimum of 4 (velars produced by the two CAS children GE and RA) to a maximum of 39 (dentals produced by an adult control subject, ME). A mixed ANOVA with 2 between factors (Subject status, Articulatory Place) and 1 within factor (Trials) has been performed on the dependent measure *k* coefficient. The levels of the Subject status factor were four: CAS subjects, lexical peers, chronological peers and adults. Since the statistical analysis did not reveal any significant interaction between the trials factor and the status factor, we excluded that CAS subjects were more variable from trial to trial than control subjects, and collapsed all the repetitions together.

We then run an ANOVA keeping only the factors Subject status and Articulatory Place as variables. Table 2 shows the number of syllables (or F2c-F2v pairs) on which the Locus Equations has been calculated. The values of the *k* coefficient of the equations (which indexes the degree of coarticulation) are reported on Figure 8, averaged across trials.

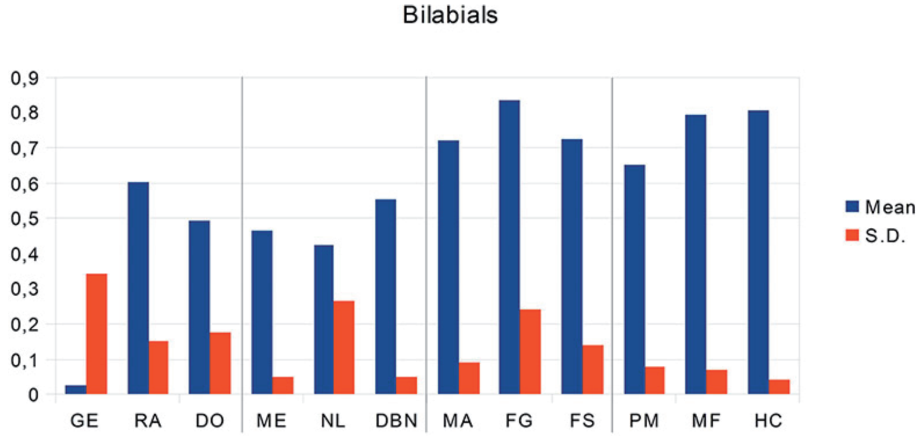
Table 2 - *Number of CV and CVC syllables selected for the coarticulation analysis in each subject, summed across the three trials*

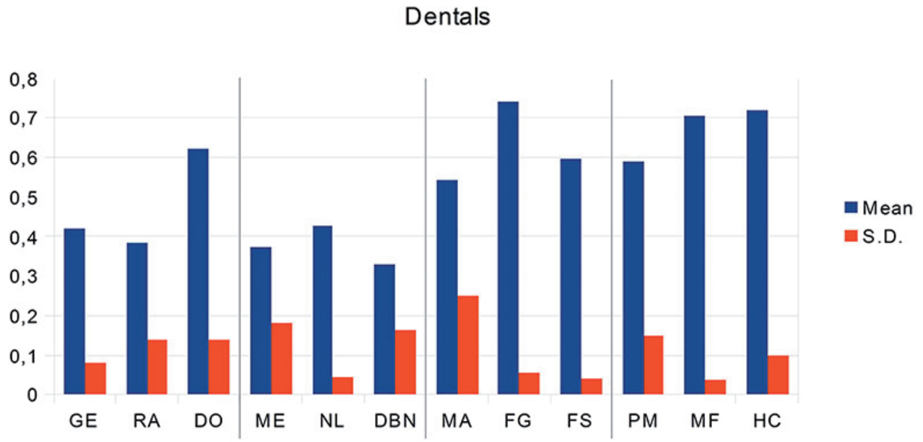
	CAS			*Lexical peers*			*Chronol. peers*			*Adults*		
	GE	RA	DO	ME	NL	DBN	MA	FG	FS	PM	MF	HC
BIL.	52	43	74	76	51	76	63	66	63	67	61	62
DEN.	73	59	79	114	78	107	69	79	79	85	71	79
VEL.	27	9	40	53	27	84	47	43	51	43	48	55

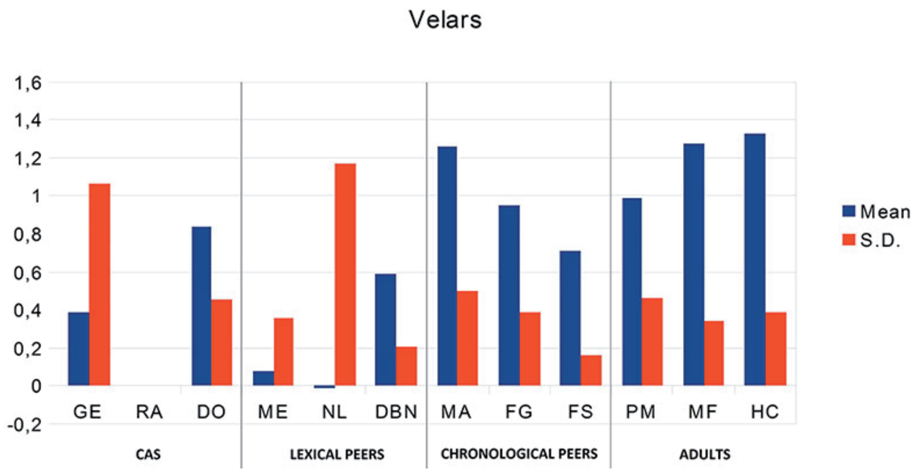
Results show that Subject status and Articulatory Place are statistically significant ($F_{(3,93)} = 13.166$, $p = 0.000$; $F_{(2,93)} = 3.419$, $p = 0.037$ respectively), as well as their interaction (F (6,93) = 2.337, p = 0.038). A post-hoc Tukey's Honestly-Significant-Difference Test on Subject status evidences that CAS and Lexical Peers have the lowest *k* values (i.e. less coarticulation: 0.486 and 0.359, respectively) and contrast significantly (p < .05) with Chronological Peers (0.786) and Adults (0.873).

Three ANOVAs, performed on the Subject status factor within each place of articulation (Bilabial, Dental, Velar) shows that CAS and Lexical Peers have the lowest *k* values and always contrast significantly (p < .05) with Chronological Peers and Adults. As a further analysis, we performed 4 mixed ANOVAs on *k* values within each Subject status condition, keeping Articulatory Place as a between factor and Trials as a within factor. Results show that only adults distinguish significantly the three places of articulation (F(2,6) = 15.730, p = 0.004).

Figure 8 - *The values of k, the coefficient of the Locus Equations indexing the degree of the anticipatory coarticulation of V on C, averaged across tree trials (S.D.: standard deviations)*







An ANOVA on Standard Deviations values resulting from the k values in the three trials does not evidence any significant difference, although CAS children always exhibit the higher SD values, indicating a trend towards more variability.

4. *Discussion*

According to the segmental analysis based on the phonetic transcription, the multiple productions (until three times) of the same speech target were phonologically more inconsistent (from trial to trial) for the two CAS subjects rather than their lexical peers, and the percentage of incorrect words increased with length and complexity of words. There were, however, exceptions: for instance, the CAS subject RA showed more correct disyllabic than monosyllabic words (see above, Footnote 4).

To sum up, CAS subject speech was characterized by widespread abnormalities at different levels, in phonological segments as well as in stress allocation. These results agree with the existing literature (e.g. Chilosi *et al.*, 2014). Whereas the performances of the two CAS subjects were quite similar to one another, the control subject ME performed much better than the control subject NL. This discrepancy might raise a question about the criteria of selection of control subjects; we will keep it as a warning for the future. However, the alternative choice of matching the subjects for the segmental analysis of the words produced by means of a naming task (i.e. the phonetic test) on the basis of the chronological age would have been even worse.

As to the analysis of intra-syllabic CV anticipatory coarticulation, a third CAS subject was added, and he was administered the phonetic test. All the CV and CVC syllables in the intelligible lexical productions elicited through the phonetic test were analysed. The three CAS subjects and the lexical age peers obtained low coefficients of coarticulation and they were significantly different from the chronological age peers and the adults, suggesting that CAS children could have an immature speech motor control (like that of 2- or 3-years-old normally developing children). Anyway, we did not find the CAS children to be significantly different from controls, as to intra-individual variability among multiple repetitions of the same lexical target, as it was initially hypothesized.

The finding of weaker coarticulation in CAS children as compared to age-matched controls (but similar to lexical peers aged from two to three years) is difficult to interpret. We still do not know if the CAS children are different from age-matched controls. If they are, then a question can be raised, about whether the developmental meaning of the difference is abnormality or delay. In other words, we need to know both the characteristics of normal speech development and the characteristics of a normal mature speech system to be able to reliably recognize immaturity. Only a few studies, using different methodologies as summarized in the introduction, have compared directly the coarticulation of CAS subjects to that of normally developing children and normal adults, and they arrived at different conclusions. Weaker coarticulation in CAS subjects with respect to age-matched peers was found by Nijland *et al.* (2002), but the coarticulation degree of the age-matched peers was stronger than in the adults. In a later experiment (Nijland *et al.*, 2003), the coarticulation pattern was found to be stronger in CAS children than in the age-matched controls, and both of them were stronger than the adult coarticulation pattern. Since the authors related greater immaturity to stronger coarticulation pattern as a possible consequence of a more global and undifferentiated planning and programming of the syllables (Nittrouer, Studdert-Kennedy & McGowan, 1989), this result could qualify CAS subjects as immature. According to this interpretation, the low degree of coarticulation exhibited by the CAS subjects of our experiment would not qualify them as immature.

The pattern of weaker coarticulation that we found in CAS children was also found by Sussman *et al.* (2000), and their discussion could bring a new light into the question. Sussman *et al.* (2000) start from the established fact that in adult speakers the regression lines, interpolating the x,y coordinates of F2 values of each syllable, have a typical slope, described by the coefficient *k* of the Locus Equation, which associates to the place of articulation of the consonant. In normal condition, the value of k is maximal for velars, intermediate for bilabials and minimal for dentals. This ordering was confirmed several times by independent researchers and has been confirmed also for Italian (Petracco, Zmarich, 2007; Zmarich *et al.* 2013). According to Sussman et al (2000: 309) “LE slopes directly encode extent of anticipatory coarticulation by capturing the strength of the vowel’s influence on the preceding stop consonant”. The CAS children of Sussman *et al.* (2000) experiment, differently than age-matched controls, did not show this clear ordering. In the words of Sussman et al (2000: 309), they “showed a reduced distinctiveness in coarticulatory extents across stop place categories as indexed by LE slopes”. The authors advance the suggestion that this lack of contrastiveness contributes to the lower intelligibility of the speech of CAS speakers. Interestingly, low *k* values and lack of contrastiveness across stop place categories were also found as a result of our experiment in typical developing children much younger than CAS, matched to CAS subjects on the basis of the same number of words attested in CDI (the so called lexical peers). This similarity could suggest that CAS children could suffer from an immature speech motor control, which is characteristic of 2- or 3-years-old children (Raccanelli, 2016). This hypothesis could find some confirmation in the fact that even at a more advanced age between 42 and 47 months, a group of typically developing children was found to have lower *k* values and lower distinctiveness across stop place categories than adults (Zmarich *et al.*, 2013). Possibly, the same reasons invoked by Zmarich *et al.* (2013) to interpret their results could find an application here: in the case of bilabials, the occlusion (being anatomically not binding for tongue) provides the adult speaker a maximal temporal overlap of the tongue and lip gestures. The degree of anticipatory coarticulation is maximal since at the time of the stop release the back of the tongue is already in position for the vowel (coarticulation as co-production). The lower *k* values of CAS subjects are possibly due to the fact that the tongue is still in motion at the moment of release, indicating that these children still have to learn to synchronize the tongue occlusion movement with the release. As to dentals, CAS children possibly still need to learn to differentiate and coordinate the tip and the back of the tongue, which can be moved quite independently one from the other even though they belong to the same organ. In the case of velars the biomechanical constraints are maximal (C and V are articulated with the back of the tongue), and in adult speech the acoustic effect highlights a very high degree of coarticulation as a reciprocal adaptation. Possibly in CAS children the consonant does not change the place of articulation as a function of the vowel's place, forcing the articulators to an extra effort to reach the required positions to produce the vowel quickly.

As a final note, some caution about the interpretation of these results could come from the absence of a further analysis regarding speaking rate, which could be an important determinant of coarticulation patterns (Agwuele, Sussman & Lindblom, 2008), especially when CAS subjects, whose speech rate is notoriously slow and laborious, are considered. Our next effort will be then that of trying to relate coarticulatory patterns to syllable duration.

Bibliography

American Speech Language Hearing Association (2007). *Childhood Apraxia of Speech: Ad Hoc Committee on Apraxia of Speech in Children*.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders (DSM-5).* Washington, DC: American Psychiatric Pub.

Agwele, A., Sussman, H.M. & Lindblom, B. (2008). The Effect of Speaking Rate on Consonant Vowel Coarticulation. In *Phonetica*, 65, 194-209.

Bearzotti, F., Fabbro, F. (2003). Test per la valutazione delle Prassie orofacciali nel bambino, Protocollo B e C. In *Giornale di Neuropsichiatria dell'Età Evolutiva*, 23, 406-417.

Bertelli, B., Bilancia, G. (2008). *VAUMeLF: Batteria per la valutazione dell'attenzione uditiva e della memoria di lavoro fonologica nell'età evolutiva*. Firenze: Giunti O.S.

Bishop, D.V.M. (2009). *TROG-2: test for reception of grammar, version 2.* Firenze: Giunti O.S.

Caselli, M.C., Pasqualetti, P. & Stefanini, S. (2007). *Parole e frasi nel "Primo Vocabolario del Bambino"*. Milano: Franco Angeli.

Chilosi, A.M., Cipriani, P. (1995). *TCGB Test di comprensione grammaticale per bambini*. Edizioni del Cerro.

Chilosi, A.M., Lorenzini, I., Cerri, B. & Cipriani, P. (2014). Disprassia Verbale Evolutiva: inquadramento clinico e diagnosi differenziale con il disturbo fonologico. In Marotta, L., Caselli, M.C., (Eds.), *Disturbi del linguaggio, caratteristiche, valutazione, trattamento*. Trento: Erickson, 145-161.

Chilosi, A.M., Lorenzini, I., Fiori, S., Graziosi, V., Rossi, G., Pasquariello, R., Cipriani, P. & Cioni, G. (2015). Behavioral and neurobiological correlates of childhood apraxia of speech in Italian children. In *Brain and Language*, 150, 177-185.

Grigos, M.I., Moss, A. & Lu, Y. (2015). Oral Articulatory Control in Childhood Apraxia of Speech. In *Journal of Speech, Language, and Hearing Research*, 58, 1103-1118.

Krull, D. (1988). Acoustic properties as predictors of perceptual responses: A study of Swedish voiced stops. In *Phonetic Experimental Research at the Institute of Linguistics*, Stockholm University, 7, 66-70.

Maassen, B., Nijland, L. & Terband, H. (2010). Developmental models of Childhood Apraxia of Speech. In Maassen, B., Van Lieshout, P.H.H.M (Eds.), *Speech motor control: New developments in basic and applied research*. Oxford: Oxford University Press, 243-258.

Nijland, L., Maassen, B., van der Meulen, S., Gabreëls, F., Kraaimaat, F.W. & Schreuder, R. (2002). Coarticulation patterns in children with developmental apraxia of speech. In *Clinical Linguistics & Phonetics*, 16, 461-483.

Nijland, L., Maassen, B., van der Meulen, S., Gabreëls, F., Kraaimaat, F.W. & Schreuder, R. (2003). Planning of syllables in children with developmental apraxia of speech. In *Clinical Linguistics & Phonetics*, 17, 1-24.

Nittrouer, S., Studdert Kennedy, M. & McGowan, R.S. (1989). The Emergence of Phonetic Segments Evidence from the Spectral Structure of Fricative-Vowel Syllables Spoken by Children and Adults. In *Journal of Speech, Language, and Hearing Research*, 32, 120-132.

Petracco, A., Zmarich, C. (2006). La quantificazione della coarticolazione nello sviluppo fonetico. In Giordani, V., Bruseghini, V., Cosi, P. (Eds.), *Scienze vocali e del linguaggio. Metodologie di valutazione e risorse linguistiche.* Torriana (RN): EDK Editore srl, 135-150.

Raccanelli, E. (2016). Un'analisi acustica della coarticolazione e del VOT in bambini con Disprassia Verbale Evolutiva (DVE). Tesi di Laurea in Logopedia, Università di Padova.

Rancan, G. (2015). Applicazione di un nuovo Test Fonetico a soggetti con Disprassia Verbale Evolutiva e a soggetti con sviluppo tipico di età tra i 36 e i 48 mesi. Tesi di laurea in Logopedia, Università di Padova.

Rose, Y., Stoel-Gammon, C. (2015). Using PhonBank and Phon in studies of phonological development and disorders. In *Clinical Linguistics & Phonetics*, 29, 686-700.

Rustioni Metz Lancaster, D. (2007). *Prove per la valutazione della comprensione linguistica*. Firenze: Giunti O.S.

Stella, G., Pizzioli, C. & Tressoldi, P.E. (2000). *PPVT Peabody – Test di vocabolario recettivo.* Torino: Omega.

Sussman, H.M., McCaffrey, H.A. & Matthews, S.A. (1991). An investigation of locus equations as a source of relational invariance for place categorization. In *The Journal of the Acoustical Society of America*, 90, 1309-1325.

Sussman, H.M., Duder, C., Dalston, E. & Cacciatore, A. (1999). An Acoustic Analysis of the Development of CV Coarticulation. A Case Study. In *Journal of Speech, Language, and Hearing Research*, 42, 1080-1096.

Sussman, H.M., Marquard, T.P. & Doyle, J. (2000). An Acoustic Analysis of Phonemic Integrity and Contrastiveness in Developmental Apraxia of Speech. In *Journal of Medical Speech-Language Pathology*, 8, 301-313.

Terband, H., Maassen, B. (2010). Speech Motor Development in Childhood Apraxia of Speech: Generating Testable Hypoteses by Neurocomputational Modeling. In *Folia Foniatrica & Logopaedica*, 62, 134-142.

Terband, H., Maassen, B., van Lieshout, P. & Nijland, L. (2010). Stability and composition of functional synergies for speech movements in children with developmental speech disorders. In *Journal of Communication Disorders*, 44, 59-74.

Velleman, S.L. (2011). Lexical and phonological development in children with childhood apraxia of speech-a commentary on Stoel-Gammon's 'Relationships between lexical and phonological development in young children'. In *Journal of Child Language*, 38, 82-86.

Williams, P., Stackhouse, J. (1998). Diadochokinetic skills: Normal and atypical performance in children aged 3-5 years. In *International Journal of Language & Communication Disorders*, 33, 481-486.

Zmarich, C., Bonifacio, S. (2005). Phonetic inventories in Italian children aged 18-27 months: a longitudinal study. In *Proceedings of INTERSPEECH 2005 – Eurospeech, 9th European Conference on Speech Communication and Technology.* NewYork (NY): ISCA, 757-760.

Zmarich, C., Fava, I., del Monego, G. & Bonifacio, S. (2012). Verso un "Test Fonetico per la Prima Infanzia". In Falcone, M., Paoloni, A. (Eds.), *La voce nelle applicazioni*. Roma: Bulzoni, 51-66.

Zmarich, C., Bortone, E., Vayra, M. & Galatà, V. (2013). La coarticolazione e il VOT nello sviluppo fonetico: studio sperimentale su bambini dai 42 ai 47 mesi d'età. In Galatà, V. (Ed.), *Multimodalità e multilingualità: la sfida più avanzata della comunicazione orale.* Roma: Bulzoni, 475-493.