Word position and stress effects in consonant cluster perception and production

Article

Accepted Version

Word position and stress effects in consonant cluster perception and production

Abstract

The aim of the present study was to investigate whether the saliency effect for word beginnings reported in children with Dyslexia (Marshall & van der Lely, 2009) can be found also in TD children. Thirty-four TD Italian children aged 8-10 completed two specifically designed tasks: a production task and a perception task. Both tasks used nonwords containing clusters consisting of plosive plus liquid (eg. pl). Clusters could be either in a stressed or in an unstressed syllable, and could be either in initial position (first syllable) or in medial position (second syllable). In the production task children were asked to repeat the non-words. In the perception task, the children were asked to discriminate between two nonwords differing in one phoneme belonging to a cluster by reporting whether two repetitions were the same or different. Results from the production task showed that children are more accurate in repeating stressed than unstressed syllables, but there was no difference with respect to position of the cluster. Results from the perception task showed that children performed more accurately when discriminating word initial contrasts than when discriminating word medial contrasts, especially if the cluster was unstressed. Implications of this finding for clinical assessments are discussed.

Background

There is substantial evidence from linguistics and psycholinguistics to suggest that word initial syllables are processed differently from word medial and word final syllables. This can be explained on the basis that word initial syllables are strong, since they license a large number of contrasts and resist reduction (Beckman, 1998, 2013; Smith, 2002, 2005; Marshall & van der Lely, 2009). Non-initial syllables, on the other hand, are weak, since they license a smaller number of contrasts and tend to reduction (Marshall & van der Lely, 2009; Harris, 2011). Word initial onsets permit a greater number of sounds than word medial onsets, and resist the application of otherwise regular alternations (Smith, 2002, Beckman, 1998, 2013). They play a crucial role in lexical access (Zwisterlood, 1989; Marslen-Wilson & Zwisterlood, 1989; Pitt & Samuel, 1995) and are more likely to be recalled in the Tip of the Tongue phenomenon (Browman, 1978). Word final material, instead, is subject to deletion (Harris, 2011) and has worse priming effects on neighbouring sounds than word initial material (Marslen-Wilson & Zwisterlood, 1989). In short, word initial positions appear to be salient compared to other positions. This phenomenon was formalised by Beckman (1998), and will be reported below as *word beginning saliency principle*.

Most research on word position effects has concentrated on the lexicon and on position effects on lexical access (Brown & McNeill, 1966; Browman, 1978; Cole (1973), Cole & Jakimik, 1980; Marslen-Wilson, 1984; Nooteboom, 1981). Only a few studies have investigated word position effects at the sublexical level (Pitt & Samuel, 1995; Marshall &van der Lely, 2009), and even fewer studies address word position effects in perception at the sublexical level (Pitt & Samuel, 1995). There is evidence from existing research on English by Marshall & van der Lely (2009) that word position effects at the sublexical level are found in clinical populations, such as children dyslexia and/or SLI. The analyses of Beckman (1998, 2013) and Smith (2002, 2005) suggest that the word beginning saliency principle is a general principle that applies to human phonology, thus one should expect word position effects at the sublexical level to be also found in the TD population, and in languages different from English, such as Italian. Furthermore, if we reconsider the word beginning saliency principle within recent models of phonology, such as Ramus et al. (2010), it seems reasonable to expect word position effects in perception as well as production. The reason is explained in detail in the following paragraph:

The word beginning saliency principle is described as a constraint within a theory of phonology known as optimality theory (Beckman, 1998). Generative and optimality theories of phonology naturally describe a unidirectional process: production (Ramus et al, 2010). The classical generative theory distinguishes between underlying and surface representations (Chomsky & Halle, 1968, 1990). Underlying representations are stored forms of words, in which some phonological traits are underspecified. Surface representations are the result of the application of phonological rules of the language on the underlying representations. Optimality theory develops this idea and distinguishes between lexicon and post-lexicon (Prince & Smolensky, 1997). The term “post-lexicon” refers to the output form of a given word, after a set of constraints has been applied to the lexicon (Beckman’s analysis (1998) belongs to this account). Based on generative models of phonology, Ramus and colleagues (2010) developed an Information Processing Model (IPM) which takes into account perception as well. The IPM (Ramus et al, 2010) proposes the existence of a lexicon and the existence of a sublexicon. The former contains prototypical word forms, while the latter contains information about the phonological rules to be applied in perception and production to map speech using these prototypical forms. Ramus et al (2010) explicitly state that their distinction between lexicon and sublexicon corresponds to the generative distinction between underlying and surface representations (Chomsky & Halle, 1968), but this distinction, as explained above, accounts only for the output pathway of their model. In order to also account for the input pathway, Ramus et al. (2010) divide sublexical representations into Output and Input Sublexical Representations. The Input Sublexical Representations level is tuned during language acquisition, and contains a mapping of the phonemes of a given language and information on relevant and irrelevant contrasts. The Output Sublexical Representations contain surface forms of words (Chomsky & Halle, 1968), i.e. phonological variations of lexical forms derived through phonological processes related to the rule of the language, the context, the register. Input and Output Sublexical Representations mutually influence, and partially shape each other, even if the relation between the two levels is poorly understood (Ramus et al, 2010). It seems evident, however, that these levels are not entirely independent from each other, and can be indistinguishable in monolingual adults (ibid). For this reason, if a principle, such as the one proposed by Beckman (1998), applies in production, it might also be expected to be found in perception. Our tests investigate the access to these two distinct levels of phonological representation.

The present study builds on work by Marshall and van der Lely (2009). In their study the authors showed that children with developmental dyslexia and/or SLI have more problems in repeating nonwords containing consonant clusters found in word medial syllable onsets than if they are in word initial syllable onsets, and children with developmental dyslexia only (no co-morbidity) are less accurate for consonant clusters in unstressed than in stressed syllables. Given the theoretical foundation of their study (particularly, given the work of Beckman, 1998, 2013 and Ramus et al, 2010), we expect similar word position effects to be found in similarly aged TD children in perception and production.

Hypothesis

We hypothesise that the word beginning saliency principle proposed by Beckman (1998, 2013) and detected in production by Marshall and Van der Lely (2009) in clinical populations is a general principle that applies to both perception and production of any spoken material in both typical and atypical children and across languages. This predicts that there will be better accuracy in the detection and production of consonant clusters in word beginnings compared to the detection and production of consonant clusters in word medial syllables in typical Italian-speaking children.

Method

Thirty-four children from a state primary school in Siena (Tuscany, Italy), aged 8;03 to 10;01, were recruited (Mean age 8;09, SD, 6 months, 19 M). None of the children had a diagnosis of developmental disorders. The children were seen individually. Children’s non verbal abilities were assessed using the Coloured Progressive Matrices (Raven, 1995). The mean standard score for the CPM was 98, sd. 15. No child scored lower than 2SDs below the mean. Individual scores are available in Appendix 1.

Reading performance was also assessed using a standardised measure of reading performance for Italian called Batteria per la Valutazione della Dislessia e della Disortografia Evolutiva - DDE-2 (Sartori et al, 2007). The children completed subtests 2 and 3. Subtest 2 is a real word reading task, consisting of 4 types of words: highly concrete and frequent words, highly concrete and infrequent words, highly abstract and frequent words, and highly abstract and infrequent words. Subtest 3 is a nonword reading task, consisting of three types of words: short shallow words, long shallow words, and opaque words generated with regular orthographic rules (for more details, see appendix 1). Considering that Italian has a shallow orthography, TD children between the ages of 8 and 10 are quite accurate in reading; hence the time needed to perform the reading task is usually preferred as a measure of variability. The results showed that children's mean reading time was 183 seconds (sd. 54). Reading accuracy was at ceiling and as a consequence we are confident in excluding the presence of phonological/reading deficits.

Production Task: The production task required the child to repeat 40 non-words containing clusters. The words used were trisyllabic and contained only the vowel /a/. Accuracy was measured. Non-words were presented in a child-friendly context. Children watched a video of a dancing parrot that seemed to pronounce the 40 non-words. They were asked to repeat what the parrot was saying as accurately as possible. The video could be stopped at any point by the child pressing the space bar and was re-started by pressing the same key. Nonwords were generated so that each contained a phonological cluster. The cluster was always formed of a plosive consonant, followed by a liquid consonant, followed by the vowel /a/. The cluster could be either in the first or in the second syllable, and stress was either in the first or in the second syllable. This gives 4 conditions:

Description of Stimulus	Example
1. cluster first syllable, stress first syllable	i.e. pla:kata
2. cluster first syllable, stress second syllable	i.e. plaka:ta
3. cluster second syllable , stress first syllable	i.e. ka:plata
4. cluster second syllable , stress second syllable	i.e. kapla:ta

Clusters were formed as a combination of plosives and liquids so that, in word medial position, the two consonants were always processed as belonging to the same syllable. According to Roach (1991, 2000) this type of cluster is never decomposed, and there is no risk of the plosive being interpreted as the coda of the previous syllable. Ten words for each condition were created. For a complete list see appendix 2.

Perception Task: The perception task contained 40 pairs of words. Half of the word pairs were identical words and half were pairs of words differing in one phoneme generating a minimal pair. Children were asked to press white when they thought the two words were identical and black when they thought the two words were different. The words used were trisyllabic and contained only the vowel /a/. When words in the pair differed in one phoneme, this phoneme was always part of the cluster, and the difference was always of one single trait: voicing. This contrast has been used in several previous studies (for a review, see Hoonorst, 2011). For instance, pairs of differing words included "tra:kata / dra:kata" or "praka:ta / braka:ta". Clusters were positioned in the first or in the second syllable, which was either stressed or unstressed. Thus there were four possible conditions in which the two words differed, and four possible conditions in which the two words were the same:

Description of stimulus	Target	Different	Same
Cluster in the first syllable, stress in the first syllable:	tra:kata	dra:kata	tra:kata
Cluster in the first syllable, stress in the second syllable	traka:ta	draka:ta	traka:ta
Cluster in the second syllable, stress in the second syllable	katra:ta	kadra:ta	katra:ta
Cluster in the second syllable, stress in the first syllable	ka:trata	ka:drata	ka:trata

Results:

Correlations:

Initially, an analysis of correlation between age and accuracy in all the tasks was performed, in order to understand if age accounts for significant variance in accuracy. None of these correlations was significant. Age and Perception Accuracy, $r = .24$, $p > .05$, Age and Production Accuracy, $r = -.25$, $p > .05$. Thus, age was not related to task accuracy and so was not considered in further analyses.

Accuracy in the perception task was found to correlate significantly with reading performance, $r = .38$, $p < .05$. Accuracy in the production task (calculated dividing the number of errors by the number of given answers) did not correlate with reading performance, $r = -.19$, $p > .05$, but partial correlation between number of missed answers and reading time (with accuracy as control) was significantly correlated with reading time using a one-tailed hypothesis (justified, for instance, by Torgesen and Burgess, 1998), $r = .28$, $p = .05$.

Production:

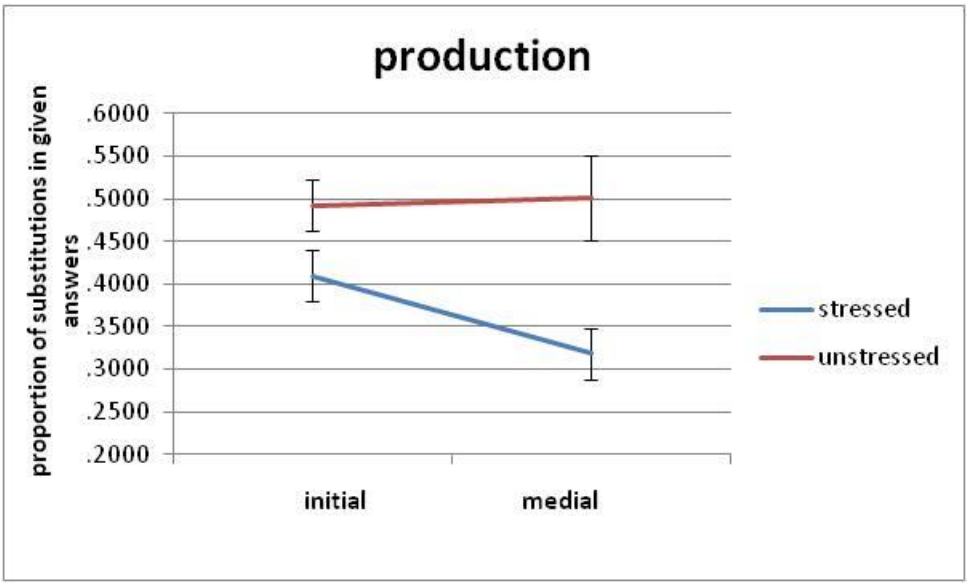

Figure 1: Production: comparison of the means in the four conditions. Children made more errors in unstressed compared to stressed syllables. Word position effects were absent. The interaction between position and stress was also marginally significant.

Next an error analysis was performed. Deletions were quite rare in this task occurring thus less than once in every hundred words (26 errors in 2720 non words presented), and were therefore not analysed separately. Instead, deletions and substitution errors were combined in one analysis. The relation between errors, stress and word position was analysed using two way ANOVA: the first factor was the position of the cluster (word initial and word medial), the second factor was whether clusters were stressed or not (cluster stressed, and cluster unstressed).

The analysis of errors shows a significant effect of stress, $F(33, 1) = 23.096$, $p < .001$, with children making more errors in unstressed compared to stressed syllables. There was no effect of word position, $F(33, 1) = 1.84$, $p > .05$, but we detected a marginally significant interaction, $F(33, 1) = 3.82$, $p = .059$. Post-hoc analysis shows that the contrast between stressed and unstressed syllables in word medial position is highly significant, $t(33) = -4.08$, $p < .001$, and that the same contrast in word initial position is only marginally significant, $t(33) = -2.45$, $p = .02$ (Bonferroni adjusted alpha = .025).

Perception:

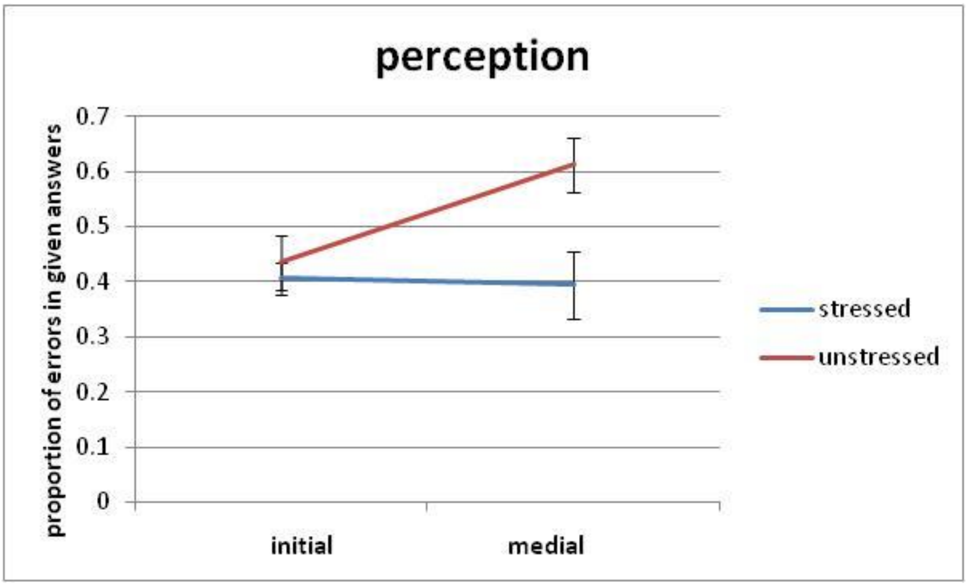

Figure 2: Perception: comparison of the means across stressed and unstressed, and initial and medial clusters. There was a significant effect of stress with children making more errors in the unstressed than the stressed condition. There was also a main effect of position with children making more errors in the medial than initial position. However, there was an interaction between stress and position: children showed no different in rate of errors between stressed and unstressed syllables in the initial condition and so differences between stressed and unstressed syllables were limited to the medial position.

Being a same-different task, we checked for the presence of biases using d-prime analysis. We calculated hit rate, false alarm rate and the d-prime value for each participant. We then compared the d-prime values to 0 and 1 using one sample t-tests. The t-tests showed that the d-prime values are significantly different from 0, indicating that performance is not random (MacMillan & Creelman, 2005), $t(33) = 8.27$, $p < .0001$, and they are also significantly bigger than 1, indicating an overall accuracy for both *different* and *same* trials of more than 70% (ibid), $t(33) = 3.42$, $p = .002$, two tailed.

In order to investigate word position and stress effects we then conducted a two way ANOVA having position of the cluster and stress as factors. The two way ANOVA shows a significant word position effect, $F(33, 1) = 12.76$, $p = .001$. Children made more errors in the detection of contrasts when the clusters were in the medial than the initial position (initial vs medial means, .421, .503, SE, .026, .036). There was also a significant effect of stress, $F(33, 1) = 14.75$, $p = .001$, with children making more errors when the clusters were unstressed than stressed (stressed vs unstressed, .400, .524, SE, .032, .032). Finally, there was a significant interaction, $F(33, 1) = 8.18$, $p = .007$. Post-hoc comparisons showed that children made significantly more errors in the medial position when the syllable was unstressed rather than stressed ($t = 4.38$, $p < 0.0001$) and made more errors in unstressed syllables when the cluster was in the medial position compared to when the cluster was in the initial position ($t = 5.67$, $p < 0.0001$). Other comparisons did not reach significance. (see figure 2).

Conclusion

We hypothesised that the word beginning saliency principle proposed by Beckman (1998, 2013) and detected in production by Marshall and Van der Lely (2009) in clinical populations is a general principle that applies to both perception and production of any spoken material in both children with typical and atypical language and/or literacy development.

The main hypothesis is confirmed for perception and production: word position effects are present in TD Italian children. In the perception task, the participants were more accurate in the discrimination of word initial contrasts than in the discrimination of word medial contrasts, if the clusters were unstressed. With regard to the production task no word position effect was detected, but the data showed a significant stress effect, with stressed clusters being repeated more accurately, and the marginally significant interaction suggests that the stress effect is driven by word medial positions (thus, indirectly, it shows a word position effect). These results extend Marshall and van der Lely's work (2009). They also confirm the word beginning saliency principle (Beckman, 1998, 2013) in perception and production, and are in line with the predictions of Ramus et al (2010)'s model of phonological representations. Further, accuracy in the perception task was found to correlate significantly with reading performance, extending to Italian a cross-linguistically well established correlation between voicing contrast perception and reading (Hoonorst, 2011), and performance in the production task was found to partially correlate with reading, adding relevant material to the debate about the relation between short-term memory and reading (Torgesen and Burgess, 1998).

These findings may have consequences on well-established assessments for children with language and/or reading difficulties such as the Children's Test of Nonword Repetition (CNRep, Gathercole & Baddeley, 1996). The CNRep assesses working memory (which often correlates with both language and reading abilities) and is often used as part of a battery in the assessment of developmental disorders (ibid.). In this test there are 4 types of nonwords, divided according to number of syllables:

10 two syllable words, 10 three syllable words, 10 four syllable words and 10 five syllable words. Normative data suggest that longer nonwords are repeated less accurately by all age groups (ibid.). However, this claim does not take into account word position effects generated by clusters. We showed in our study that non initial clusters are processed less accurately than word initial clusters. Inspection of the distribution of non-initial clusters in the CNRep task shows that they are not balanced across syllable length. Non-initial clusters are all positioned in the four- and five syllable words, and never in the two- and three- syllable words. A chi square shows that the distribution of clusters in non-initial position is significantly unbalanced: χ (3) = 11.9, p = .008 two tailed. This suggests that the normative data obtained for the CNRep assessment may be influenced by the unbalanced distribution of non-initial clusters, not only by the length of the word.

In conclusion, in this paper we report evidence that word position and stress effects affect the way children perceive and produce nonwords, with word beginnings being perceptually salient. This finding should be taken into account when using non-word tasks in the assessment of children with language and/or reading difficulties.

Appendices

Appendix 1

Reading test:

Orthographic productive rules used:

“giu” = /dʒu/
“sce” = /ʃe/
“gn” = /ɲ/
“gli” = /ʎi/

Real words, highly concrete and frequent: i.e. vino (wine), bambino (child), letto (bed)
Real words, highly concrete and infrequent: i.e. insetto (bug), cero (wax), margine (edge)
Real words, highly abstract and frequent: i.e. pace (peace), ragione (reason), successo (success)
Real words, highly abstract and infrequent: i.e. dominio (domination), sciopero (strike), simbolo (symbol)
Nonwords, shallow and short: i.e. fosto, prisi, tonca
Nonwords, shallow and long: i.e. locostato, tacipaca
Nonwords, opaque: gnoba, pronounced ɲɔba, cogiu, pronounced cɔdʒu

Coloured Progressive Matrices standardised scores

Id	Age	Score	Standardised score

s1s1	~	34	~
s2s1	9;0	29	108
s3s1	9;3	30	100
s4s1	8;10	31	115
s5s1	9;3	30	100
s6s1	9;9	29	100
s7s1	9;2	26	90
s8s1	8;10	19	70
s9s1	9;2	34	120
s10s1	8;11	21	75
s11s1	9;0	31	105
s12s1	8;7	28	100
s13s1	10;0	22	75
s14s1	9;6	25	85
s15s1	9;1	31	105
s16s1	9;9	32	110
s17s1	~	25	~
s18s1	8;6	34	130
s19s1	8;2	26	90
s20s1	8;3	26	90
s21s1	9;2	32	110
s22s1	9;11	29	100
s23s1	9;3	30	100
s24s1	8;4	27	95
s25s1	8;4	26	90
s26s1	8;5	28	100
s27s1	8;6	29	105
s28s1	8;3	33	125
s29s1	9;0	32	110
s30s1	8;6	30	110
s31s1	8;10	21	75
s32s1	8;9	27	95
s33s1	8;7	25	85
s34s1	8;3	23	80

Appendix 2

Specifically designed stimuli: non-words

Cl1 str1 unvoiced	**Cl1 str1 voiced**
tra:kata	dra:kata
pla:kata	bla:kata
pra:kata	bra:kata
kla:kata	gla:kata
kra:kata	gra:kata
Cl1 str2 unvoiced	**Cl1 str2 voiced**
traka:ta	draka:ta
plaka:ta	blaka:ta
praka:ta	braka:ta
klaka:ta	glaka:ta
kraka:ta	graka:ta
Cl2 str2 unvoiced	**Cl2 str2 voiced**
katra:ta	kadra:ta
kapla:ta	kabla:ta
kapra:ta	kabra:ta
kakla:ta	kagla:ta
kakra:ta	kagra:ta
Cl2 str1 unvoiced	**Cl2 str1 voiced**
ka:trata	ka:drata
ka:plata	ka:blata
ka:prata	ka:brata
ka:klata	ka:glata
ka:krata	ka:grata

References:

Beckman, J. N. (1998, 2013). Positional Faithfulness: an Optimality Theoretic Treatment of Phonological Asymmetries, New York, Routledge.

Browman, C. P. (1978). Tip of the tongue and slip of the ear: Implications for language processing, University of California, *UCLA Working Papers in Phonology*, 42, 1-159.

Brown, R., & McNeill, D. (1966). The “tip of the tongue” phenomenon. *Journal of verbal learning and verbal behavior*, 5(4), 325-337

Chomsky, N., & Halle, M. (1968). The Sound Pattern of English, New York. Harper and Row.

Cole, R. A., & Jakimik, J. (1980). How are syllables used to recognize words? *The Journal of the Acoustical Society of America*, 67(3), 965-970

Gathercole S. E. & Baddeley A. L. (1996), The Children’s Test of Nonword Repetition (Manual), The Psychological Corporation.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, 2 (2), 103-127.

Harris, J. (2011) Deletion. In: van Oostendorp, M and Ewen, CJ and Hume, EV and Rice, K, (eds.) The Blackwell Companion to Phonology. (1597 - 1621). Wiley-Blackwell: Oxford, UK.

Hoonhorst, I., Medina, V., Colin, C., Markessis, E., Radeau, M., Deltenre, P., & Serniclaes, W. (2011). Categorical perception of voicing, colors and facial expressions: A developmental study. *Speech communication*, 53(3), 417-430.

Macmillan N. A. , Creelman C. D. (2005), Detection theory. A user’s guide. Psychology press. Taylor and Francis group. New York, London.

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology*, 15(3), 576.

Nooteboom, S. G. (1981). Lexical retrieval from fragments of spoken words: beginnings vs. endings. *Journal of Phonetics*, 9(4), 407-424

Pitt, M. A., & Samuel, A. G. (1995). Lexical and sublexical feedback in auditory word recognition. *Cognitive Psychology*, 29(2), 149-188

Prince, A., & Smolensky, P. (1997). Optimality: From neural networks to universal grammar. *Science*, 275 (5306), 1604-1610.

Ramus F. , Peperkamp S. , Christophe A. , Jacquemot C. , Kouider S. , Dupoux E. (2010), A psycholinguistic perspective on the acquisition of phonology, *Laboratory phonology* (10), 311-340.

Raven J. C. (1995), Coloured Progressive Matrices, Pearson Assessment, San Antonio, Texas.

Roach P. (1991, 2000), English Phonetics and Phonology, Cambridge.

Marshall C. R. , Van der Lely H. K. J. (2009), Effects of word position and stress on onset cluster production: evidence from typical development, SLI and Dyslexia, *Language*, 85, 39-57.

Sartori G. , Job R. , Tressoldi P. E. (2006), DDE-2, Batteria per la Valutazione della Dislessia e della Disortografia Evolutiva, Giunti O.S.

Smith, J. (2002, 2005). Phonological Augmentation in Prominent Positions. New York, Routledge.

Torgesen, J. K., & Burgess, S. R. (1998). Consistency of reading-related phonological processes throughout early childhood: Evidence from longitudinal-correlational and instructional settings. In J. L. Metsala & L. C. Ehri, (Eds.), Word recognition in beginning literacy (pp. 161–188). Mahwah, NJ: Erlbaum.

Zlatin, M. A., & Koenigsknecht, R. A. (1975). Development of the voicing contrast: Perception of stop consonants. *Journal of Speech and Hearing Research*, 18(3), 541.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, 32(1), 25-64.