Name: Alisha Hixson

Student ID: 1703482

Supervisor: Dr Lawyer

Project Title: The development of the perception of native allophones in children aged 6- 15-years old.

Pages: 60

Word count*: 8641

Not including tables/titles

Contents

List of chapters

List of tables:

List of graphs:

List of figures:

Introduction

Fikkert (1994, p.14) defines perception as “discrimination and identification”. The literature suggest that perception differs depending on whether the contrasts are phonemic or allophonic. Phonemic contrasts occur when the two sounds can be used to form different words (Hayes, 2009), such as the difference between /bɒks/ *(box)* and /fɒks/ *(fox)*. The literature suggests that phoneme perception increases with age, although it is not clear how this changes over 10-years-of-age. On the other hand, allophonic contrasts do not form different words (Hayes, 2009). For example, [anθm̩] and [an̪θm̩] are different pronunciations, but are not different words. Allophone perception ability appears to become language specific early in development but trends in adulthood are less clear.

This project aimed to investigate some of these gaps in the literature by testing the perception of native phonemes and allophones in children aged 6-15-years old and an adult control group using three tasks; AX discrimination, identification, and a rating task, based on the method in Iverson et al. (2003).

As well as mapping changes in development, the project considered two factors which the literature suggests may influence perception. The first was place of articulation. This is because the “Motor Theory” of perception (Liberman & Mattingly, 1985, p.2) and the “Perceptual Assimilation Model” (“PAM”) (Best & McRoberts, 2003, p.183) suggest that “articulatory gestures” (Jusczyk, Shea & Aslin, 1984, p.453), may be the units of perception. Therefore, differences in place of articulation may influence perception. To research this, the perception of [l-ɫ] and [n-n̪] were compared.

Secondly, the study considered the influence of accent on perception as studies suggest experience may influence perception (Iverson et al., 2003; Kelly, 2017). To investigate this, the perception of an accent feature, glottalisation, was tested with the contrast [t-ˀt].

Literature Review

Phoneme perception

Several studies have been conducted on the development of native phoneme perception, finding that as age increases, perception ability also increases, a trend which has been found with stimuli with varying manners of articulation. This shows a phenomena named "categorical perception" which Altmann (2001, p.25) explains as when perception is stronger for sounds that belong to different phoneme categories, such as /d/ and /g/, than sounds which belong within a category, such as two variants of /b/ with different voice onset times.

Many of these studies have been conducted with plosives. These suggest that the ability to discriminate these contrasts is present from a young age. Jusczyk and Thompson (1978), who used the high amplitude sucking method, found that 2-month-old infants were able to discriminate /g/ and /b/ in both the initial and medial positions of CVCV structures. This finding was also reached by Eimas (1974) using the same method. Eimas (1974) suggests that this ability is linguistic; when the stimuli were presented as non-speech sounds, the infants failed to show categorical perception. Other methods have been used to test the perception of this contrast in infants; Moffit (1971) measured the heart rate of 4-5-month old infants to investigate the perception of /b/ and /g/. This supported the findings of the previous research; responses implied that the infants were able to discriminate the contrast.

This ability appears to increase with age. Ohde and Haley (1997) researched the perception of /b/, /d/ and /g/ in 3-4-year olds compared to an adult control group. Adult participants were asked to press a button to respond whereas the child participants were asked to indicate which of two puppets' names they heard, for example, "/bi/" (Ohde & Haley, 1997, p.3712). There was a significant effect of age on performance in the discrimination for all of the contrasts, suggesting that there was a difference in perception between 3-4-year olds and adults.

It is not clear when the development of phoneme perception reaches adult-like levels. Sussman and Carney (1989) used an AAXX discrimination task on a synthetic /b/-/d/ contrast where participants were asked to indicate whether there was a change between the stimuli. This found that their 10-year-old participants did not perform at adult-like levels which implies that perception abilities are still developing after 10 years of age. However, the child participants were less accurate as they tended to incorrectly report a change in stimuli when the contrasts were the same, suggesting that, rather than being unable to discriminate a contrast, they incorrectly reported that there was one present. Furthermore, the participants performed similarly in terms of labelling. The adult data suggests that labelling corresponded with discrimination, so if sounds were perceived as different, the participants gave the stimuli different labels. In contrast, the child participants performed similarly to the adults in identification but not discrimination, so they were able to identify stimuli that they appeared to discriminate inaccurately. Therefore, it is not clear from this study how perception changes from the age of 10-years-old to adulthood although, prior to this age, perception ability appears to increase.

Increasing perception ability has also been reported for approximants; using the high amplitude sucking method, Jusczyk, Copan and Thompson (1978) tested 2-month old infants' perception of native /w/ and /j/ contrasts and found that the infants appeared to be able to discriminate the pair. These findings have been replicated with other contrasts; Eimas (1975) presented 2-3-month-old English acquiring infants with stimuli from two phonemic categories, /l/-/r/, or from within one of the categories, /l/-/l/ or /r/-/r/. As with the plosives, discrimination was stronger between than within the categories, implying that the infants were able to discriminate the contrast.

This ability is retained in adulthood for speakers who use the phonemes contrastively in their native language. Iverson et al. (2003) tested German and English speakers on the /r/-/l/ contrast as well as Japanese learners, for whom the acquisition of the /r/-/l/ contrast is challenging. Participants were given a range of tasks including identification, a rating task, and an AX discrimination task (Iverson et al., 2003). The results found that German and English speakers were able to discriminate the /r/-/l/ contrast, whereas the Japanese learners did not show increased sensitivity at the /r/-/l/ boundary. Iverson et al. (2003) also noted that Japanese learners tended to identify the stimuli as /r/ rather than /l/. Therefore, the participants perceived the contrasts more accurately when they were found in the native language.

Similar research has been conducted with nasals. Narayan, Werker and Beddor (2010) tested 4-to-5-month-old, 6-8-month old and 10-12-month old infants on their perception of

/n/ and /m/ using the visual habituation method. The study found longer looking times when the participants were played different stimuli than when the same stimuli was played. This suggests that the infants were able to discriminate between the two consonants. However, when tested using the same method, the infants were unable to reliably discriminate the /n/-/ŋ/ contrast. The researchers note that this pair is more "acoustically similar" than the /m/-/n/ contrast which may have been why the /n/-/ŋ/ pair was more difficult for the infants to discriminate (Narayan, Werker & Beddor, 2010, p.407).

However, by adulthood, participants show the ability to discriminate this contrast. Larkey, Wald and Strange (1978) presented participants with stimuli containing /m/, /n/ and /ŋ/ in an AXB discrimination task, where participants were asked to identify which of three sounds was the odd sound out, and found that the participants were able to discriminate between the three nasal categories.

So, although the findings vary slightly depending on place of articulation, the literature suggests that phoneme perception ability is present from a young age and increases for contrasts which are found in the native language. However, it is not clear how phoneme perception changes over the age of 10-years-old.

Allophone perception

As with phonemic perception, the literature suggests that in the early stages of development, infants are able to discriminate allophonic contrasts; Hohne and Juscyzk (1994) conducted research with 6-12-week old infants using the high amplitude sucking method. They found that the infants were able to discriminate different variants of /tr/ where the first was aspirated and retroflex [ʈʰ] with a voiceless [ɹ̥] and the second an unreleased [t̚] and voiced [ɹ]. This ability appears to be retained until 4-months old as Seidl, Cristia, Bernard and Onishi (2009), using the head turn preference method, found that 4-month-old English infants were able to discriminate oral-nasal vowel allophone contrasts, like those found in French. However, by 11-months-old, the infants showed a bias for contrasts that are found in their native language; when Seidl et al. (2009) tested 11-month-old infants, they found that the French infants had retained the ability to discriminate the contrast but English infants of the same age were not able to do so.

This trend has been reported in adult speakers. Harnsberger (2001) conducted research on the discrimination of nasal contrasts with speakers from a range of languages using an AXB discrimination task. One of the participant groups were English speakers who were able to discriminate nasal contrasts with phonemic status such as /am:a/-/an:a/ (99-100%) (Harnsberger, 2001, p.496). However, they were less accurate in non-phonemic trials, such as when discriminating /an̪a/-/aɳa/ (44-89%) (Harnsberger, 2001, p.496). This has been replicated with plosives; Whalen, Best and Irwin (1997) tested American English speakers on an AXB discrimination task and found that participants were more accurate at discriminating /b-p^h/ and /b-p/ than the allophonic contrast /p-p^h/.

As well as being influenced by the native language, the literature suggests that the variety of the language spoken may influence perception; Kelly (2017) researched perception with American and Irish English speakers with a dental-alveolar contrast; /t/ and /d/ in contrast to /t̪/ and /d̪/. This is contrastive for Irish speakers who may produce dental fricatives as stops. The study found that the Irish speakers were significantly more accurate in discriminating the contrast than the American English speakers, for whom the stimuli were not contrastive.

However, other studies have been conducted which have not found this decline in perception ability. Johnson and Babel (2010) presented Dutch and English participants with paired stimuli including combinations of /s/, /θ/ and /ʃ/. They found that Dutch speakers rated /s/ and /ʃ/, which in Dutch is an allophone of /s/, as more similar than the English speakers, who use /s/ and /ʃ/ contrastively. This supports the proposal that allophone perception decreases with age. However, the different participant groups performed similarly in a speeded AX discrimination task. This may be due to the research task used; the study used a speeded AX discrimination task, which has a "low memory load" (Johnson and Babel, 2010, p.132) whereas the other allophone studies used a standard discrimination or identification task. Therefore, the difference in task may have influenced the results.

Similarly, Peperkamp, Pettinato and Dupoux (2013) found that allophone perception ability depended on the structure of their stimuli. Their French participants were tested on /ʁ/ and /χ/, which are allophonic in French, compared to /m/ and /n/ which are phonemic. Participants were presented the stimuli either in "isolation" in VC syllables or in "context", in a VCCV structure (Peperkamp, Pettinato and Dupoux, 2003, p.152). The study found that in the context condition, participants had difficulty in discriminating the allophones although they were able to discriminate the phoneme contrast. On the other hand, when they were presented with the stimuli in isolation, although there was still a significant difference between the phoneme and allophone condition, participants made significantly fewer errors.

So, although the literature suggests that infants are able to discriminate allophonic contrasts it is not clear whether this declines into adulthood.

Theories of perception

The literature offers several theories of perception. The two relevant for this project are the "Motor Theory" (Liberman & Mattingly, 1985, p.2) and the "Perceptual Assimilation Model" ("PAM") (Best & McRoberts, 2003, p.183).

According to the Motor Theory, "articulatory gestures" (Jusczyk, Shea & Aslin, 1984, p.453), such as tongue and lip movements, are the units used in perception (Liberman & Mattingly, 1985). So, sounds which are produced by similar gestures are perceived as similar or the same whereas sounds which are produced by different gestures remain discriminable (Jusczyk, Shea & Aslin, 1984). These units are made of a class of movements rather than a single movement, meaning that, although phonemes are be produced differently in different contexts, they are perceived as the same unit of sound (Liberman & Mattingly, 1985).

This is supported by neurological evidence. D'Ausilio et al. (2009) presented participants with bilabial and apical plosives to discriminate but, before the stimuli was presented, administered TMS pulses to cortical regions associated with the lips or tongue. The researchers found, as predicted, that stimulation resulted in faster reaction times in discriminating the sounds. This supports the Motor Theory as it implies that motor regions may be involved in speech perception.

This is also supported by recent behavioural research with infants by Choi, Bruderer, Werker (2019). In this study, 5-month-old infants were given teething toys which limited either tongue or lip movement in a head turn preference study. The infants were first tested on bilabial, alveolar and retroflex contrasts without the toy, which found that they were able to discriminate the contrasts. However, when they were presented with a teething toy which prevented tongue movement, there was no significant difference in looking time between the alveolar and retroflex stops, suggesting that the infants did not discriminate the contrast. Likewise, when a teething toy preventing lip closure was given to the infants, they were no significant differences in looking times between bilabial and alveolar stops, again suggesting that infants did not discriminate the contrast. This suggest that, in line with the Motor Theory of perception, “oral-motor processes” may influence perception (Choi, Bruderer, Werker, 2019, p.1396).

Best (1991) addresses the Motor Theory, suggesting that infants are firstly able to perceive all speech but that by the end of the first year of life, infants have received enough input to detected gestural patterns in input. Then, by around 4-years of age, children are able to detect which speech sounds are better examples than others and by the time adulthood is reached, perception occurs phonemically, and “unfamiliar” sounds are grouped with native speech sounds based on their gestural quality (Best, 1991, p.21).

This led to the Perceptual Assimilation Model (PAM) which suggests that, depending on the differences in the articulatory gesture, non-native speech sounds are perceived in different ways (Best & McRoberts, 2003). Firstly, there are “single-category” contrasts where, if two sounds are similar to one native phoneme, they will be difficult to discriminate (Best & McRoberts, 2003, p.186). These are more difficult to discriminate than “category-goodness” pairs which occur when two sounds are attributed to the same native category but it is recognised that one is a better fit than the other (Best & McRoberts, 2003, p.186). The easiest contrasts to discriminate, according to the model, are “two-category” pairs which occur when each non-native phoneme in a pair is attributed to two different native contrast, (Best & McRoberts, 2003, p.186). Whalen, Best and Irwin (1997) suggest that allophone contrasts are perceived as category goodness when the differences are detectable or as single category contrasts when they are difficult to discriminate.

Like the Motor Theory, the PAM has been evaluated with research evidence. Harnsberger (2001) predicted the responses in the AXB discrimination tasks based on the predictions made by the PAM. This reported that perception for category-goodness and single category contrasts was more accurate than the model predicted and also that many of the stimuli pairs were unable to be categorised according to the model. Nonetheless, the analysis found that the two-category contrasts were best discriminated (97%), then category goodness (75%) then single-category (67%), which is the order predicted by the model (Harnsberger, 2001, p.497). So, although this study suggests that there are some limitations to the PAM, it does lend some support to its predictions.

Aims

Based on gaps in the literature, this project aimed to investigate the development of perception in children between the ages of 6-15-years old in order to research changes in perception over the age of 10-years-old and to investigate the development of allophone perception due to the varying results in the literature. This may have implications for speech and language therapy. This is because the research may provide a suggested outline of development which could be contrasted to the perception of children and adults diagnosed with language impairments, such as dyslexia which is associated with impaired perception abilities (eg. Serniclaes, van Heghe, Mousty, Carré & Sprenger-Charolles, 2004), in further research.

In order to investigate the development of perception and the influence of speech gestures and accent, the following research questions were formed:

1. Does perception change between 6-15-years-old?
2. Does development occur in stages?
3. Does place of articulation influence perception?
4. Does accent influence perception?

The following hypotheses were made:

1. As age increases, perception ability and the difference in perception between the control and experimental stimuli will increase.
2. Perception may develop in stages.
3. As articulatory gestures may influence perception (e.g. Best, 1991; Liberman & Mattingly, 1985), stimuli pairs with a larger difference in place of articulation will be more easily discriminated. Therefore, [l-ɫ] will be easier to discriminate than [n-n̪].
4. The [t- ˀt] contrast may be the most easily discriminated. This is because glottalisation is a feature of the East England accent (Trudgill, 1974), where the study was conducted and is described as a “shibboleth” by Chambers (2009, p.203). Based on Kelly (2017), as this is a feature of their accent, participants may be more accurate in discriminating the contrast.

Method

Participants

A total of 87 participants were recruited for the project. The number of participants by age and gender is shown in Table 1:

Age group	Average age (yrs.months)	N =	Male	Female
6-years old	6.07	6	3	3
7-years old	7.05	6	3	3
8-years old	8.05	5	3	2
9-years old	9.6	8	3	5
10-years-old	10.05	9	4	5
11-years-old	11.06	7	2	5
12-years-old	12.06	8	5	3
13-years-old	13.06	13	7	6
14-years-old	14.06	11	4	7
15-years-old	15.05	4	2	2
Adult	19.11	10	5	5
TOTAL		=87	=41	=46

Table 1: Participants by age and gender

The following criteria were applied in recruiting participants:

- No history of hearing impairments.
- No history of language impairments. This is because research suggests that there is a relationship between language impairment and atypical perception abilities (eg. Serniclaes et al., 2004).
- Monolingual English speakers. This is because the literature suggests that language experience can affect perception abilities (eg. Iverson et al. 2003).

The child participants were recruited through opportunity sampling and through a volunteer sample in schools in the Ipswich area. The four schools gave the children participant information sheets and consent forms for the parent/carers to complete and return to the school if consent was given. For children in primary school, written consent was gained from parent/carers and for high school students, written consent was gained from both the parent/carers and the students. In addition, before the experiment began, participants gave verbal consent to take part in the research. Participants were tested either at school or at home. The adult participants were recruited through an opportunity sample, gave written and verbal consent and were tested at their work place or homes. On the consent form, participants provided their age and gender and adult participants and the parent/carers of the child participants were asked to state where they grew up to control for the effect of accent. After recruitment, participants were assigned a reference number to ensure anonymity. This consisted of their school year group or ‘C’ for adults and a letter from ‘A’ to ‘J’.

Stimuli

The stimuli were 12 VCV syllables read by a female native English speaker. Six stimuli were part of the phoneme control condition and six formed the experimental allophone condition forming three pairs for each condition as shown in Table 2.

	Individual Stimuli		**Paired Stimuli**	
	Control	**Experimental**	**Control**	**Experimental**
Nasal	/ama/	[ana]	/ama-ama/	[ana-ana]
			/aŋa-aŋa/	[an̪a-an̪a]
	/aŋa/	[an̪a]	/ama-aŋa/	[ana-an̪a]
			/aŋa-ama/	[an̪a-ana]
Approximant	/aja/	[ala]	/aja-aja/	[ala-ala]
			/aɹa-aɹa/	[ała-ała]
	/aɹa/	[ała]	/aja- aɹa/	[ala-ała]
			/aɹa-aja/	[ała-ala]
Plosive	/ada/	[ata]	/ada-ada/	[ata-ata]
			/aga-aga/	[aˀta-aˀta]
	/aga/	[aˀta]	/ada-aga/	[ata-aˀta]
			/aga-ada/	[aˀta- ata]

Table 2: Individual and paired stimuli

The nasal and approximant contrasts were used to investigate the effect of place of articulation. This is because the former contrasts between an alveolar and dental place of articulation, both using the tongue front (Zsiga, 2013). Therefore, the gestures for both allophones are similar. In contrast, [l] is an alveolar lateral approximant, formed with the tongue tip and alveolar ridge whereas [ɫ] is a velar lateral approximant, formed using the tongue body and soft palate (Zsiga, 2013). Therefore, as these involve different articulators, there is a greater difference in articulatory gesture than the in nasal contrast. The plosive allophones were used to research the effect of accent as glottalisation is a feature of the accent of the region the research was conducted in (Trudgill, 1974).

Five of each stimulus were recorded on a recording app and one of each stimulus was selected and edited using Praat in the following selection process:

- Any recordings with background noise were discarded.
- The remaining recordings were played with their paired stimuli (eg. /ama/ and /aŋa/), to find the pairs with the most similar vowel quality and length to minimise differences between the recordings besides the consonant contrast.
- Recordings were cut and edited to an average of 0.542 seconds and were saved as individual stimuli.
- To form the paired stimuli for the AX discrimination task, sounds were pasted into the same Praat file with a 1 second pause inserted between each stimulus.

Procedure

Participants were given three tasks, based on the methodology in Iverson et al. (2003). The control condition was completed before the experimental condition in one sitting, with the exception of two child participants who were tested over two days as the experiment could not be completed before the end of the testing session. Instructions were given from the script shown in Appendix A and the stimuli were played through headphones so that the experimenter was unable to hear the stimuli to reduce potential experiment effects. Participants' responses were recorded on the sheet in Appendix B. Most of the primary-school aged children played a game of Connect4 during the experiment. The three tasks were:

1. *AX discrimination task*

This task tested participants' discrimination of the two sounds in a pair. The pairs were either the same (AA) or different (AB). This discrimination task was chosen rather than alternatives as, according to Locke (1980), AX discrimination is less influenced by memory than other tasks like ABX. Each possible pairing was presented five times to test for consistency between responses. Participants responded with the labels 'same' or 'different' although some used ‘yes’ or ‘one’ for same and ‘no’ or ‘two’ for different. Their responses were coded as ‘A’ when they reported the pair as the same and ‘B’ when different and later as correct or incorrect.

2. *Identification task*

This task aimed to investigate whether participants associated the stimuli with the same orthographic category. Participants were played individual stimuli and were asked to identify the sound from the letter chart in Appendix C. The chart contained 18 boxes and was based on the Jolly Phonic reading system (Jolly Phonics, n.d.) as this contains more symbols than the English orthographic system. Moreover, it was assumed that the younger participants would be familiar with the system as this is used in schools to teach reading. Each stimulus was repeated three times to test for consistency between results. The consonant indicated by the participant was noted and their responses later coded as ‘correct’ or ‘incorrect’.

In pilot trials, the chart contained 24 options but this made the task more difficult. When asked, a participant reported knowing which symbol they wanted but struggling to find the symbol on the chart. To make the task easier, some of the options, which did not correspond to the stimuli, were removed. The stimuli were also repeated five times in the pilot trials but this was repetitive and, as responses were consistent, was reduced to three trials.

3. *Rating task*

This task addressed whether participants perceived the stimuli differently when presented individually and tested category-goodness. Participants were presented with the same individual stimuli as in the identification task and were asked to rate how ‘normal’ the speech sound was. For primary school participants, this was on a dichotomous scale as ‘good’ or ‘bad’ and for the older participants, a Likert scale from 1-5 where 1 was ‘strange’ and 5 ‘normal’. The rating given by the participants was noted. This was the number for the older participants and either ‘G’ (‘good’) or ‘B’ (‘bad’) for the child participant. ‘B’ was later coded as 1 and ‘G’ as 5.

Like the identification task, the pilot study repeated each stimulus five times but due to participant boredom this was reduced to three repetitions.

Order of presentation

To prevent order effects, the stimuli were presented randomly in one of five ‘forms’ (playlists). Each form contained random orders for each task, made using Excel, which were then analysed to ensure that the same stimuli did not occur in the same environment within a task (for example, that /ama/-/aŋa/ did not occur before /ada/-/aga/ more than once). The forms were also compared to ensure that, across the forms, each task began and ended with the different stimuli. Five forms (A-E) were created as the study intended to test five males and five females in each age group.

Repetitions

Repetitions were not offered at the beginning of the test but if a sound file did not play fully or the participant requested a repetition, the stimulus was repeated. However, one participant’s data was removed from the study due to potential demand characteristics; the participant made repeated requests for repetition in the experimental condition of Task 1 which was usually followed by a rating of ‘different’, suggesting that they were aware they were being tested on a difficult different contrast.

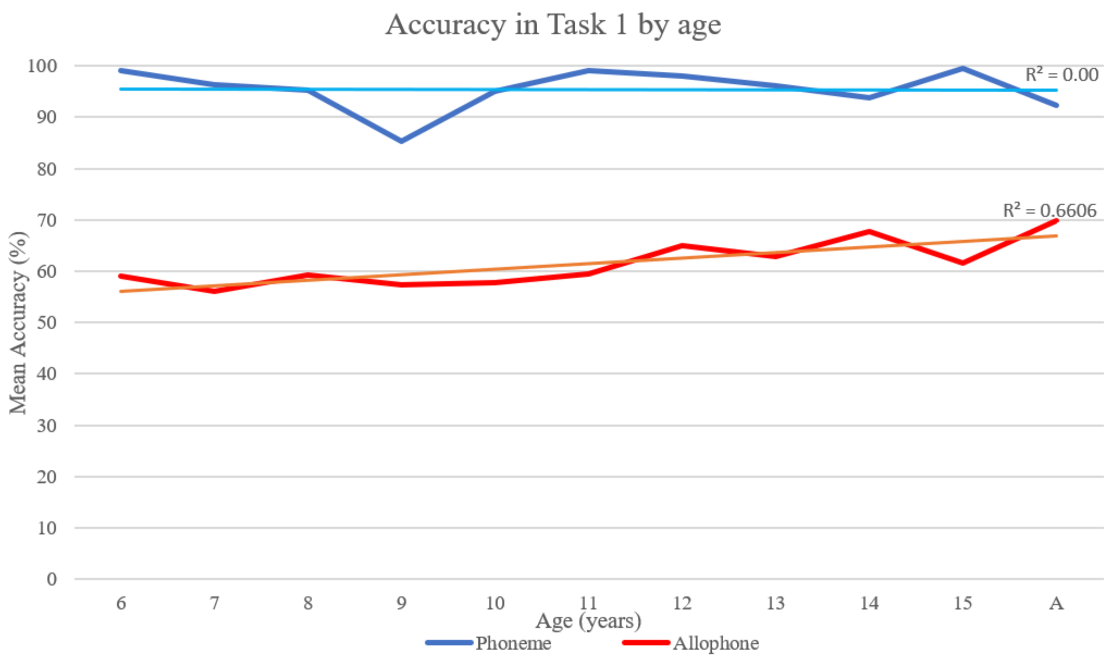
Data Analysis

Statistical models were used to analysis the results of the experiment, firstly addressing trends in age, then differences between the different stimuli.

Task 1: AX Discrimination

Age

Analysis of the data suggests that accuracy in the AX discrimination task increased with age, as shown in Graph 1.



Graph 1: Accuracy in Task 1 by age.

A binomial regression was used to test the significance of this trend by comparing the results of the child participants and adult control group, including stimuli code, gender and form as additional factors (See Appendix D for SPSS output tables). The model was significant, $\chi^2(29) = 6089.951$, $p<.01$, predicting 69.0% (Nagelkerke R^2) of the variance in accuracy and correctly classified 92.3% of cases. The difference between the adult control group and child participant group was significant ($p<.01$), which suggests that there was an increase in accuracy between child and adult participants. Stimuli code was also significant, as was Form D was significant ($p=.01$), although gender did not reach significance.

Another binomial regression was run on the child participant data, considering age, stimuli code, gender and form (See Appendix E for SPSS output tables). As with the first analysis, the model was significant, $\chi^2(29) = 5303.545$, $p<.01$. This explained 70.7% (Nagelkerke R^2) of the variance in accuracy and correctly classified 92.3% of cases. Age was significant, ($p<.01$), so accuracy increased significantly with age. As with the first analysis, stimuli code was significant. In this analysis, gender and Form D were both significant to $p<.05$.

Responses were also analysed for accuracy between AA (same stimuli) and AB (different stimuli) trial types. Table 3 shows the errors made in this trial by each age group. The boxes highlighted in red show the incorrect responses for each condition. The analysis found that the errors made in the phoneme condition between the age groups were similar with high levels of accuracy. Aside from the 10-year old participants, in the AA trials for the phoneme condition, participants were also generally accurate. Participants made the most errors in the AB trials for the allophone condition, reporting the pairs as the same, when they were different. As found by the analyses above, the number of errors decreased as age increased although there were fluctuations in this trend as shown in Graph 1.

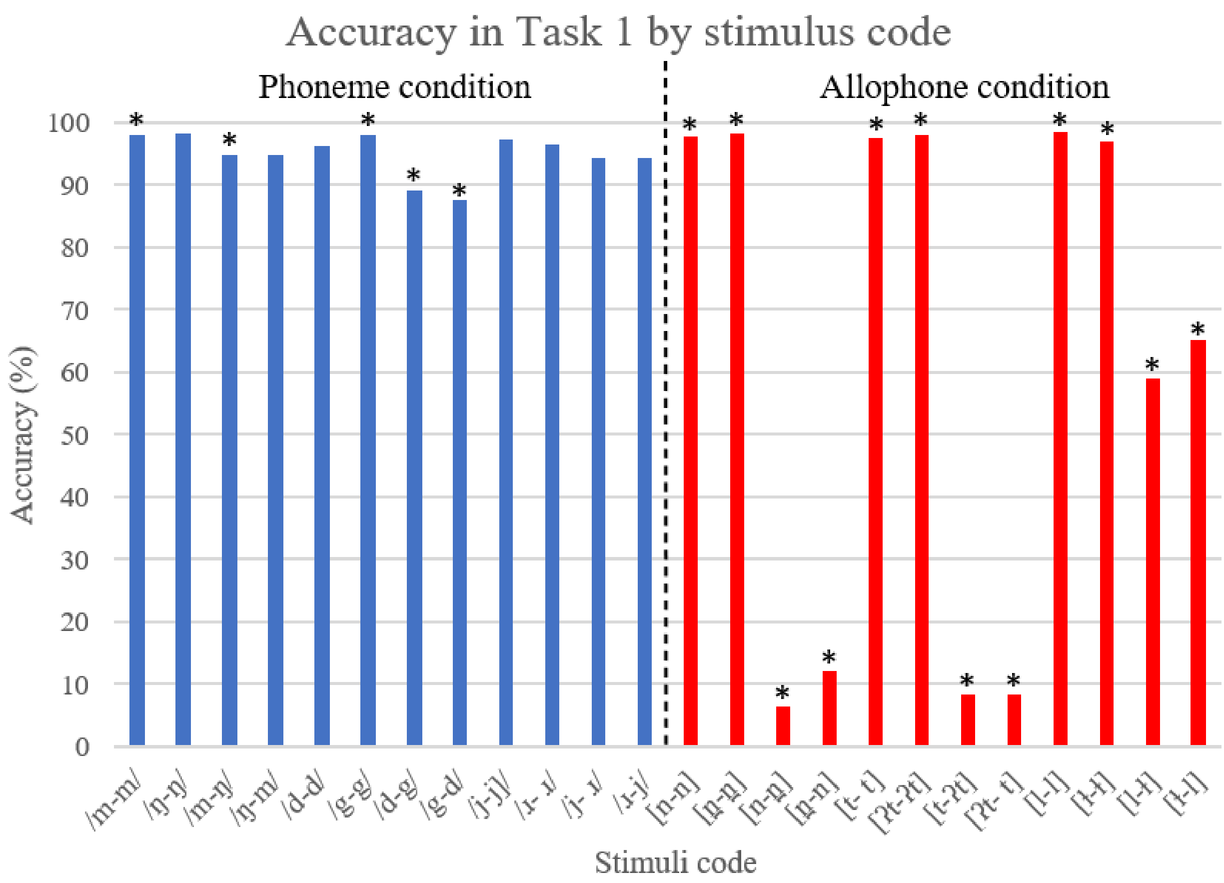
	Phoneme condition				**Allophone condition**			
	AA		**AB**		**AA**		**AB**	
Age (years)	A	B	A	B	A	B	A	B
6	28.8	1.2	0.6	29.4	28.6	1.4	23.2	6.7
7	29.3	0.7	1.5	28.5	29.83	0.17	26.16	3.84
8	29	1	1.8	28.2	29.6	0.4	26	4
9	26.75	3.25	6	24	29.25	0.75	24.875	5.125
10	28.77	1.23	1.78	28.22	27.22	2.78	22	8
11	29.71	0.19	0.29	29.71	29.71	0.29	24	6
12	29.25	0.75	0.62	29.38	29.75	0.25	20.75	9.25
13	29.77	0.23	2.15	27.85	29.77	0.23	22.3	7.7
14	29.36	0.64	3.1	26.9	29.9	0.1	20	10
15	30	0	0.25	29.75	29.25	0.75	22.25	7.75
Adults	29.3	0.7	3.2	26.8	29.5	0.5	16.6	13.4

Table 3: Errors made in Task 1 by age

Stimuli

Analysis showed that there was an effect of condition; a binomial regression was run on all of the data, considering stimuli category, participant group, gender and form (See Appendix F for SPSS output tables). The model was significant, $\chi^2(7) = 1825.684$, $p<.01$ and explained 25.1% (Nagelkerke R^2) of variance in accuracy. 78.% of cases were correctly classified. Stimuli category was significant ($p<.01$); participants were more accurate in the phoneme condition (95% correct) than the allophone condition (68% correct). The significant difference between the adult control group and child participant group was maintained ($p<.05$) although gender and form were not significant.

The different pairs were discriminated to difference levels of accuracy as shown in Graph 2.



Graph 2: Accuracy in Task 1 by stimulus code

As mentioned, the effect of stimuli code was significant. Table 4 presents condensed data from the two age regressions, showing whether the stimuli significantly differ from /m-m/ and the direction of the change:

Phoneme condition					Allophone condition				
Stimuli	Child data		All data		Stimuli	Child data		All data	
	Sig.	B	Sig.	B		Sig.	B	Sig.	B
[m-m]	0.00*		0.00*		[n-n]	0.06	0.63	0.04*	0.82
[m-ŋ]	0.02*	-0.93	0.03*	-0.88	[n-n̪]	0.00*	-5.86	0.00*	-6.06
[ŋ-m]	0.12	-0.46	0.28	-0.34	[n̪-n]	0.00*	-4.76	0.00*	-5.08
[ŋ-ŋ]	0.06	0.74	0.08	0.68	[n̪-n̪]	0.00*	1.68	0.00*	1.53
[d-d]	0.59	-0.16	0.47	-0.21	[t- t]	0.00*	1.11	0.00*	1.17
[d-g]	0.00*	-1.29	0.00*	-1.16	[t-ˀt]	0.00*	-5.02	0.00*	-5.17
[g-d]	0.00*	-1.21	0.00*	-1.02	[ˀt- t]	0.00*	-4.75	0.00*	-4.92
[g-g]	0.02*	0.85	0.03*	0.77	[ˀt-ˀt]	0.00*	1.78	0.00*	1.65
[j-j]	0.14	0.45	0.23	0.37	[l-l]	0.00*	2.11	0.00*	1.98
[j- ɹ]	0.16	-0.32	0.90	0.03	[l-ɫ]	0.00*	-1.90	0.00*	-2.09
[ɹ-j]	0.20	-0.29	0.55	0.17	[ɫ-l]	0.00*	-1.56	0.00*	-1.73
[ɹ- ɹ]	0.41	0.23	0.08	0.61	[ɫ-ɫ]	0.00*	1.37	0.00*	1.26

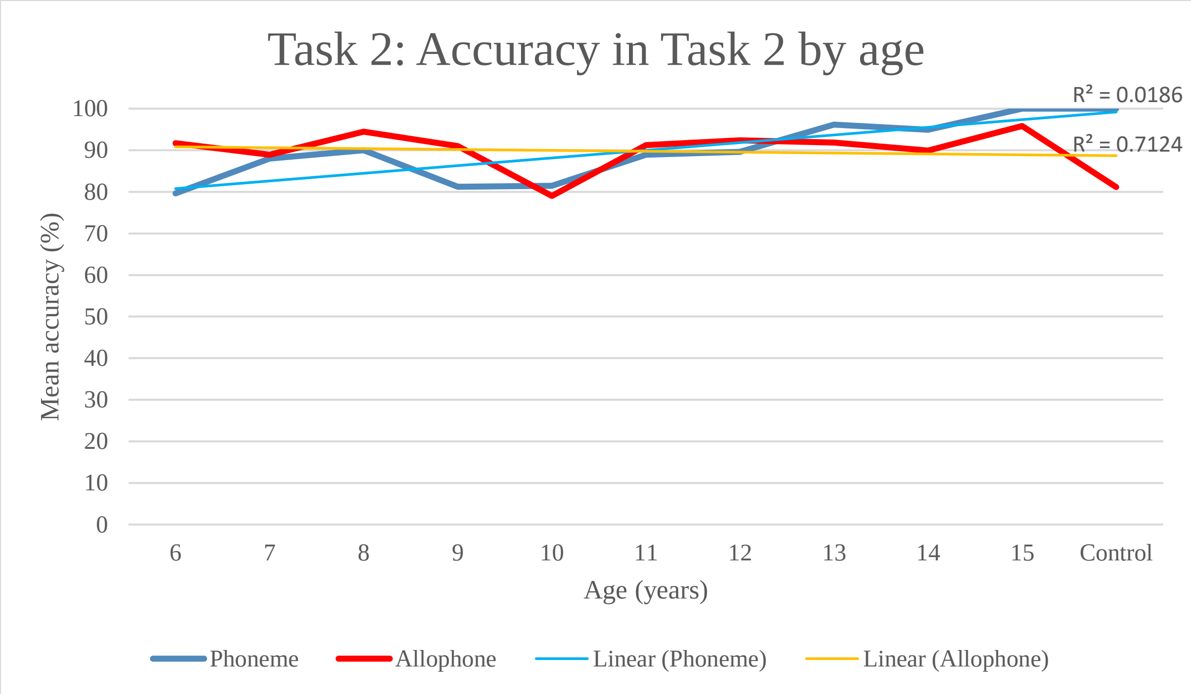
Table 4: Significance of stimuli pairs in Task 1

In the phoneme condition, participants were significantly less accurate in discriminating the plosive and one of the nasal contrasts in AB trials. However, they were significantly better at discriminating the /g-g/ contrast. All of the contrasts in the allophone condition significantly differed from the /m-m/ contrast. For the AA pairs there was a significant increase in accuracy whereas for the AB contrasts, there was a significant decrease in accuracy. This was smallest for the approximant contrast, followed by the plosive contrast then the nasal contrast. These results will be discussed in the Discussion section.

Task 2: Identification

Age

As in Task 1, there was a positive trend of age and accuracy in this task, as shown in Graph 3.



Graph 3: Accuracy in Task 2 by age

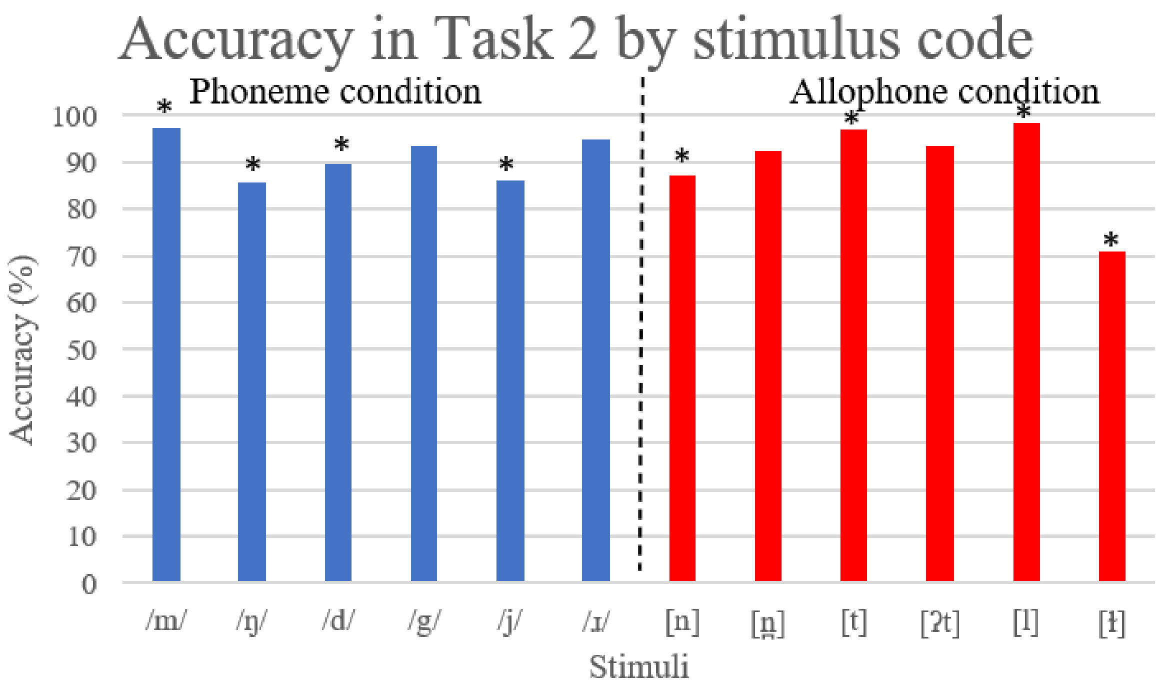
However, unlike Task 1, there was no significant difference between the adult and child participants. A binomial regression was run to test this, although there were some issues with this model (See Appendix G for SPSS output tables). The model considered the effect of participant group, stimuli code, gender and form and was significant: $\chi^2(17)= 1966.018$, $p<.01$ explaining 12.8% (Nagelkerke R^2) of variance in accuracy. The model correctly classified 90.2% of cases. Participant group was not significant ($p=.140$), suggesting that the adult control group were not significantly more accurate than the child participants. Gender also lacked significance although Form A ($p<.01$) and Form D ($p<.01$) and stimuli code were significant.

However, age had an effect in the child participant data. To test this, another binomial regression was run on age, stimuli code, gender and form on the child data (see Appendix H for SPSS output tables). The model was significant: $\chi^2(17)= 175.529$, $p<.01$, explained 13.2% (Nagelkerke R^2) of the variance in accuracy and correctly classified 90.5% of cases. Age was significant ($p<.01$), suggesting that the accuracy increased with age for the child participants. As with the previous analysis, gender was not significant but Form A ($p<.01$), Form B ($p<.05$), Form C ($p<.05$), Form D ($p<.01$) and stimuli code were significant. So, although accuracy did increase with age for the child participants, there was no significant difference in perception between the adult and child participants.

Stimuli

To test the difference between the phoneme and allophone condition a binomial regression was used, considering participant category, stimuli category, gender and form (see Appendix I for SPSS output tables). There were also issues this this model, which was significant $\chi^2(7) = 16.619$, $p<.05$, explaining 1.1% (Nagelkerke R^2) of the variance in accuracy and correctly classified 90.2% of cases. The effect of stimuli category was not significant ($p>.05$) suggesting that there was no difference in accuracy between the conditions which was 90.6% correct in the phoneme condition compared to 89.8% for the allophone condition. Gender and participant category also lacked significance although Form A ($p<.01$) and D ($p<.01$) were significant.

However, there were some differences in accuracy between the individual stimuli, as shown in Graph 4.



Graph 4: Accuracy in Task 2 by stimulus code

Table 5 shows the significance of these differences in accuracy for each stimuli code for the regression on the child data and whole data set. The analysis compares the accuracy of the stimuli to /m/.

Phoneme condition					Allophone condition				
Stimuli	Child data		All data		Stimuli	Child data		All data	
	Sig.	B	Sig.	B		Sig.	B	Sig.	B
/m/	0.00*		0.00*		[n]	0.086	-0.395	0.01*	-0.544
/ŋ/	0.00*	-1.766	0.00*	-1.799	[n̪]	0.09	0.513	0.609	0.128
/d/	0.005*	-0.813	0.001*	-0.899	[t]	0.01*	0.961	0.003*	1.08
/g/	0.183	0.423	0.361	0.269	[ˀt]	0.508	0.188	0.528	0.168
/j/	0.001*	-0.752	0.004*	-0.635	[l]	0.002*	1.55	0.001*	1.654
/ɹ/	0.077	0.547	0.045*	0.614	[ɫ]	0.00*	-1.558	0.00*	-1.789

Table 5: Significance of stimuli pairs in Task 2

For the phoneme condition, participants were significantly less accurate when identifying /ŋ/, /d/ and /j/ compared to /m/. In one of the analyses, /ɹ/ also reached significance. For the allophone condition, participants were significantly worse at identifying [ɫ] and [n]. However, they were significantly better at identifying [t] and [l].

A qualitative analysis was run on the types of errors, as shown in Table 6.

	Phoneme condition											
	/m/		/ŋ/		/d/		/g/		/j/		/r/	
	<n>	1	<r>	1	<l>	1	<ch>	1	<w>	1	<n>	1
	<h>	1	<h>	1	<p>	2	<y>	1	<g>	1	<h>	1
	<r>	1	<p>	2	<j>	2	<m>	1	<l>	2	<l>	1
	<y>	1	<w>	2	<r>	2	<ck>	2	<n>	2	<p>	1
	<w>	3	/j/	3	<b>	10	<d>	2	<h>	3	<m>	2
			<y>	3	<g>	15	<ng>	9	<ng>	3	<y>	3
			<g>	10					<r>	6	<t>	3
			<n>	21					<j>	16		
Total (/261)	7 (2.4%)		43 (16.5%)		32 (12.3%)		16 (6.1%)		34 (13.0%)		12 (4.6%)	

	Allophone condition											
	[n]		[n̪]		[t]		[ˀt]		[l]		[ɫ]	
	<w>	1	<w>	1	<w>	1	<n>	1	<p>	1	<t>	1
	<g>	1	<t>	2	<ck>	1	<ch>	1	<y>	1	<ng>	1
	<j>	1	<m>	5	<th>	5	<d>	1			<h>	1
	<h>	1	<ng>	10			<p>	2			<ch>	1
	<y>	1					<l>	2			<j>	2
	<l>	1					<ck>	2			<r>	2
	<m>	11					<th>	9			<y>	3
	<ng>	17									<w>	67
Total (/261)	34 (13.0%)		18 (6.9%)		7 (2.7%)		18 (6.9%)		2 (0.8%)		78 (29.9%)	

Table 6: Error made in Task 2 by stimuli

These results will be addressed in more detail in the Discussion.

Task 3: Rating Task

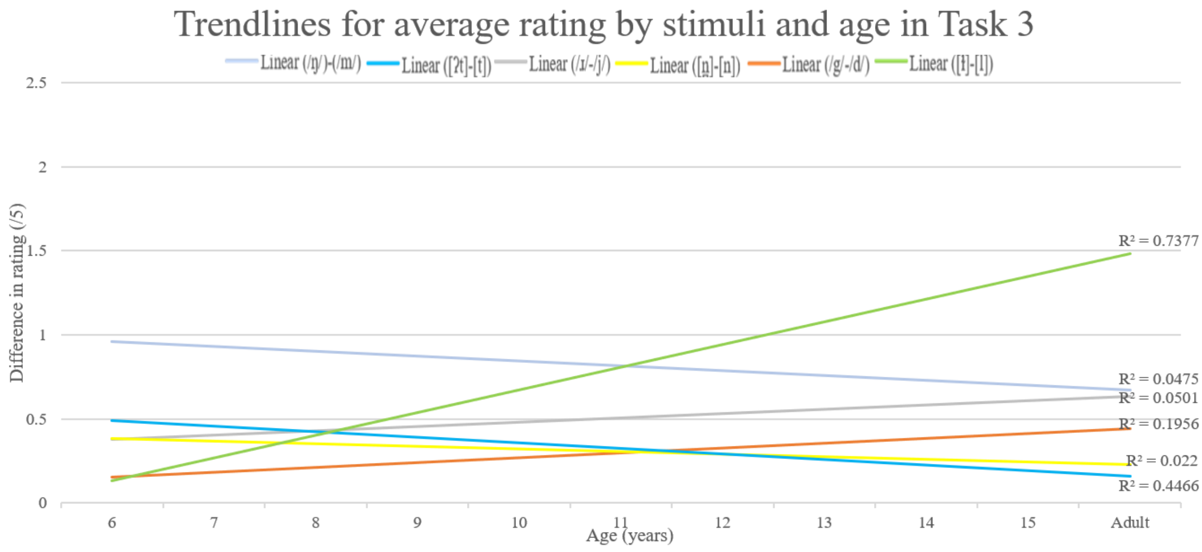
Age

The responses given by the participants varied by age, as shown in Table 7 which presents the average rating per phoneme given by each age group, with the differences between each contrast.

Phoneme condition									
Age	**/m/**	**/ŋ/**	**Difference**	**/d/**	**/g/**	**Difference**	**/j/**	**/ɹ/**	**Difference**
6	3.93	2.60	1.33	2.60	2.87	0.27	1.80	2.33	0.53
7	2.78	2.11	0.67	3.00	3.00	0	2.56	3.00	0.44
8	3.93	2.87	1.07	3.93	3.67	0.27	3.67	3.93	0.27
9	3.17	2.17	1.00	3.00	2.50	0.5	2.17	3.00	0.83
10	3.37	2.33	1.04	3.22	3.22	0	2.48	2.19	0.3
11	3.14	3.24	0.10	3.43	3.38	0.05	3.05	3.43	0.38
12	3.42	2.75	0.67	3.71	3.17	0.54	3.42	3.00	0.42
13	3.67	3.62	0.05	4.56	4.05	0.51	3.92	4.15	0.23
14	4.12	3.06	1.06	4.21	3.91	0.3	3.79	3.94	0.15
15	3.83	2.50	1.33	3.67	3.42	0.25	2.58	4.08	1.5
Adult	3.80	3.17	0.63	4.17	3.60	0.57	3.53	4.07	0.53
Allophone condition									
Age	**[n]**	**[n̪]**	**Difference**	**[t]**	**[ˀt]**	**Difference**	**[l]**	**[ɫ]**	**Difference**
6	3.40	3.40	0	2.07	2.60	0.53	2.07	2.33	0.27
7	3.00	3.00	0	2.56	3.00	0.44	2.78	2.78	0
8	3.93	2.87	1.07	3.40	3.67	0.27	2.07	2.33	0.27
9	3.33	2.67	0.67	2.83	2.33	0.5	2.17	1.83	0.33
10	3.67	3.37	0.3	2.63	2.19	0.44	2.78	2.63	0.15
11	3.48	3.48	0	3.05	2.86	0.19	2.90	2.67	0.24
12	3.79	3.25	0.54	3.46	3.00	0.46	2.79	1.79	1
13	3.77	3.62	0.15	3.82	3.62	0.21	3.33	2.49	0.85
14	4.27	4.06	0.21	4.15	3.94	0.21	3.18	1.85	1.33
15	4.25	3.83	0.42	4.42	4.42	0	3.58	2.08	1.5
Adult	4.17	4.17	0	4.37	4.07	0.3	3.93	1.87	2.07

Table 7: Average scores by stimuli and age in Task 3

These results are shown in the trendlines in Graph 5.



Graph 5: Trendlines for average rating by stimuli and age in Task 3

The difference between ratings increased with age for /d-g/, /j-ɹ/ and [l-ɫ] but decreased for /m-ŋ/, [n-n̪] and [t-ˀt]. This suggests that there were different trends with age for each stimulus.

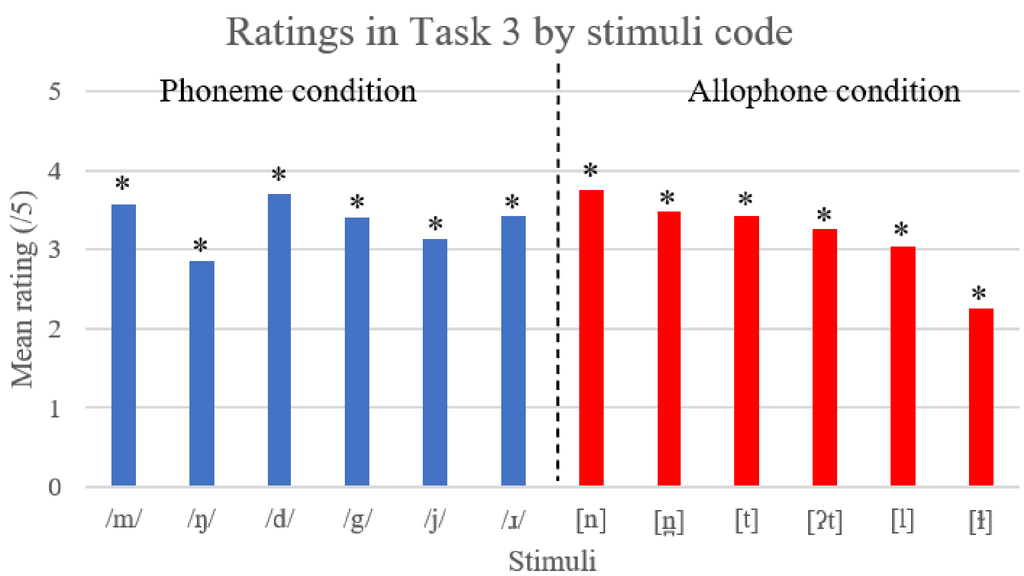
The effect of age on these ratings was significant; a multinomial regression was used to predict the participants' response to investigate the difference in ratings given by between the adult control group and child participants (See Appendix J for SPSS output tables). The model considered participant category, gender, form and stimuli code. The model included all of the data and was significant $\chi^2(17) = 252.105$, $p<.01$. The observed difference for the responses given by the different participant groups was found to be statistically significant when controlling for stimuli code, gender and form ($p<.01$). This suggests that the adults gave significantly higher ratings compared to the child participants. There was also a significant effect of gender ($p<.01$) and form when participants were tested using Form C ($p<.01$).

The second multinomial regression was used to predict the child participants' responses. This considered age, form, stimuli code and gender (See Appendix K for SPSS output tables). The model included all of the data and was statistically significant, $\chi^2(17) = 252.570$, $p<.01$. The observed difference for the responses given by the participants was found to be statistically significant for age when controlling for stimuli code, gender and form ($p<.01$). This suggests that as age increased, the rating given by the participants also increased, supporting the first model. As with the previous model, stimuli code reached significance as did Form A (p<.01), B (p<.01) and D (p<.01) although gender was no longer significant.

Stimuli

A multinomial regression was also used to predict the difference in rating between the phoneme and allophone condition. The model considered gender, form, participant category and stimuli category (See Appendix L for SPSS output tables). The model included all of the data and was significant: $\chi^2(7) = 79.3489$, p<.01. The observed difference for the responses given for each condition was significant, p<.05, when controlling for gender, form and participant category so the ratings given for the phoneme condition (3.35 out of 5 on average) was significantly higher than the average rating for the allophone condition (3.20 out of 5). The effect of gender was significant in this model (p=.01) as was form for Form C (p<.01). Participant category was also significant, as above.

The average ratings for each stimulus are shown in Graph 6:



Graph 6: Ratings in Task 3 by stimulus code

The difference between the ratings given for the nasal and plosive contrasts was smaller in the allophone condition than the phoneme condition. In contrast, the difference between the approximant pairings was greater in the allophone condition than the phoneme condition.

The previous regressions used to test the effect of age showed that stimuli code was significant. Table 8 shows the probability that within the models used to determine the significance of age and participant category, the null hypothesis for each stimulus is zero considering the presence of the other predictors in the model, and the direction of the change, compared to [ɫ].

Phoneme condition					Allophone condition				
Stimuli	Child data		All data		Stimuli	Child data		All data	
	Sig.	B	Sig.	B		Sig.	B	Sig.	B
/m/	0.00*	0.813	0.00*	0.879	[n]	0.00*	0.908	0.00*	0.984
/ŋ/	0.003*	0.315	0.00*	0.386	[n̪]	0.00*	0.701	0.00*	0.798
/d/	0.00*	0.922	0.00*	1.018	[t]	0.00*	0.623	0.00*	0.751
/g/	0.00*	0.695	0.00*	0.763	[ˀt]	0.00*	0.529	0.00*	0.644
/j/	0.00*	0.484	0.00*	0.558	[l]	0.00*	0.387	0.00*	0.512
/ɹ/	0.00*	0.664	0.00*	0.766	[ɫ]	.	0	.	0

Table 8: Significance of stimuli pairs in Task 3

All of the stimuli reached significance. These findings will be addressed in the Discussion section below.

Discussion

Does perception change between 6-15-years-old?

The hypothesis for this research question was that, as age increased, perception ability would increase and, furthermore, that the difference in the perception between the control and experimental stimuli will increase.

Task 1 supports the first part of this hypothesis; in both the phoneme and allophone condition there was a significant effect of age and a significant difference in accuracy between the adult and child participants. This suggests that, as predicted, perception increased with age. This supports the general trend in the phoneme perception literature and also the findings in Johnson and Babel's (2010) and Peperkamp, Pettinato and Dupoux' (2013) allophone perception studies. However, this may be due to the similarities in research method; the stimuli were presented in isolation, as in Peperkamp, Pettinato and Dupoux (2013), and the task was an AX discrimination task, which, although not speeded like in Johnson and Babel (2010), may be less demanding in terms of memory than alternative tasks used in the literature (Locke, 1980).

There are different possible explanations of why perception changed with age in this task. Based on Iverson et al. (2003) and Kelly (2017), discrimination ability may have been influenced by experience with the native language; adults have more experience with the contrasts, which may have a positive effect on perception, therefore as age increased, accuracy increased. Alternatively, the difference could be due to memory; Swets (1964, cited in Sussman & Carney, 1989), implies that children could be less accurate in perception tasks due to their memory processes. So, although this task is reported to be less influenced by memory than other task, it may be that accuracy increased with age due to changes in memory processing with age rather than perceptual abilities.

There was also a significant effect of age in Task 2 in the child participant data, so as age increased as did accuracy in labelling. This difference could be linked to literacy. The younger participants were relatively new to literacy so may have been less accurate whereas the older participants' literacy would be similar to adult's which may be why there was no significant difference between adults and children.

Although this research did not replicate Sussman and Carney's (1989) significant difference between adult and child participants in the identification task or trend that child participants tended to report differences in consistent stimuli, the study did find that child participants were able to identify sounds that they were unable to discriminate. Until the age of 11-years old, according to the trend lines in Graph 3, the child participants were more accurate in identifying allophones than phonemes but they were more accurate in discriminating the phoneme stimuli. Past the age of 11-years-old, the trend in identification changed; participants were more accurate in identifying phonemes rather than allophones with perception remaining stronger than the phoneme than allophone condition.

Task 3 also supports a change in perception ability although with a different trend. The differences in ratings between the stimuli pairs appear to correspond loosely to the results in the discrimination task; the difference in rating increased for [l-ɫ] and /j-ɹ/ which were the most accurately discriminated within their condition. On the other hand, the difference in ratings decreased for [n-n̪] and [t-ˀt], which were both difficult to discriminate. Therefore, it may be that the difference in rating changes with age to more closely correspond to discrimination abilities. However, this explanation cannot account for why /d-g/, which participants had some difficulty with, increased in the difference between ratings whereas the difference in ratings for /m-ŋ/, which was also significantly more difficult to discriminate in the /m-ŋ/ trial, decreased with age.

Despite this, the data supports the first part of the hypothesis for this research question; perception does change between the ages of 6-15-years old, increasing as predicted.

On the other hand, the research does not support the second hypothesis which predicted that the difference in accuracy between the perception of phonemes and allophones would increase. In the discrimination task, the opposite trend is found; phoneme perception, as in the identification task, stayed fairly stable but accuracy in the allophone condition increased. In Task 2, the difference between the conditions decreased with age until 11-years-old before increasing again, although the difference between conditions was not significant. Likewise, in Task 3, the graph does not appear to show an increase in the gap between the phoneme and allophone conditions as age increases although the difference in conditions was significant. Therefore, the difference between the conditions did not increase with age.

Does development occur in stages?

It was predicted that perception would develop in stages. This hypothesis was not supported by the data as, overall, the data in the three tasks appears to show a linear trend in perception abilities from 6-15-years old.

Nonetheless, there may be change in the progress of development from before and after 11-years-old. In Task 2, after the age of 11-years-old, phoneme perception exceeded that of allophone perception whereas before this age, the opposite trend is found, although the difference between the conditions was not significant. However, analysis from Task 3 also supports a change in perception around this age; overall, the data shows a decline in the difference in ratings until the age of 10-11-years-old, before increasing again. Therefore, although perception does not appear to occur in stages, it may be that, around 11-years-old age 11, there is some change in perception abilities.

This could be linked to the critical period as Snow and Hoefnagel-Höhle (1978, p.1114) interpret Lenneberg's (1967, cited in Snow & Hoefnagel-Höhle, 1978) critical period as ending "at about the age of puberty" due to changes in brain structure at that age. However, there are limitations with this explanation as other ages have been proposed for the critical period, such as 9-years-old (Penfield and Roberts, 1959, cited in Singleton, 2005) or younger, at 12 months (Ruben, 1997).

Does place of articulation influence perception?

The study aimed to interpret the findings in terms of the Motor Theory and perceptual assimilation model. Based on these models, it was predicted that stimuli pairs with a larger difference in place of articulation will be more easily discriminated. Therefore, the prediction was that the [n-n̪] would be more difficult than the [l-ɫ] contrast to discriminate.

The analysis supports this hypothesis; in Task 1 while only one of the control nasal contrasts reached significance (/m-ŋ/), both of the nasal allophone contrasts were significant, showing a greater decline in B value in the allophone condition than the phoneme condition. In contrast, in Task 1, the [l-ɫ] contrast showed significance but the control, /j-ɹ/ did not. However, the B value for the change in the discrimination of the approximant contrast was smaller than that of the nasal, suggesting that, although perception ability was lower, the nasal contrast was harder to discriminate.

In Task 2, there was some difference in the perception of [n-n̪]; participants made 13.0% errors for [n] compared to 6.9% for [n̪]. This suggests that, for some participants, the phonemes were associated with different categories. However, a larger difference was found for the [l-ɫ] contrast. There were 0.8% errors made for [l], which showed a significant increase in accuracy, compared to 29.9% for [ɫ], which was significantly less accurately identified, frequently as <w>.

Similar patterns between the two contrasts were found in Task 3 as both [n̪] and [ɫ] were rated lower than [n] or [l]. However, in comparison to their controls, the nasal contrasts were rated higher than the control and the approximants were rated lower. Furthermore, the difference between the ratings given the approximant allophones was greater than the difference between the nasal allophones. This suggests that the two contrasts were perceived differently.

The findings from Task 2 and 3 can be used to attribute the pairs to one of PAM's categories. Although Whalen, Best and Irwin (1997) suggest that allophone contrasts are perceived as category goodness or as single category, as [ɫ] was attributed to different categories in Task 2 and rated lower than the other stimuli in Task 3, it may be a two-category or category-goodness contrast. On the other hand, fewer errors were made with [n̪] in the identification task, suggesting that it was more consistently attributed to one category, and the ratings were higher in Task 3, suggesting that this was a single-category contrast. When these classifications are compared to Task 1, they support the predictions of PAM; [l-ɫ] was easier to discriminate [n-n̪], supporting the model's prediction that two-category and category-goodness contrasts are easier to discriminate than single category contrasts.

So, this data appears to support the prediction that difference in place of articulation influences perception. However, despite support for the use of gestures as units of perception, these findings could also be attributed to acoustic differences between the stimuli as suggested by Narayan, Werker and Beddor (2010) in their study on infant perception. Therefore, as Galantucci, Fowler and Turvey (2006) note, as gestures produce acoustic signals, evidence is required that perception occurs based on gestures rather than acoustic cues. So, although the findings can be interpreted in support of the Motor Theory, it should be noted that other explanations are possible.

Does participant awareness influence perception?

The prediction was made that the [t- ˀt] would be the best discriminated pair as it is a feature of the accent of the region the participants were recruited from. This is because Iverson et al. (2003) and Kelly (2017) suggest that accent influences perception. The results did not support this hypothesis.

Firstly, in Task 1 participants showed difficulty for the AB trials for both the control and allophone plosive contrasts with a greater decrease in accuracy in the allophone condition than the phoneme condition. Comparison of the B values in the allophone condition found that this decrease was greater than that of the approximant contrast but not of the nasal contrast. This suggests that, while the plosive contrast was more accurately discriminated than the nasal contrast, it was less accurately discriminated than the approximant pairing.

Likewise, in Task 2 participants were accurate in identifying both [t] and [ˀt], suggesting that both contrasts were strongly associated with the same category, <t>. Accuracy in these allophones was higher than that of the other stimuli in the allophone condition, suggesting that these pairing had the strongest association with the same orthographic category. Furthermore, the errors made with these allophones were similar, for example, both were identified as <ck> and <th>. Due to the strength of this category association, this suggests that the plosive contrast perceived as a single-category pairing, according to PAM, which may be why it was difficult to perceive.

Finally, in Task 3, the average rating for the allophones was lower than that of the phoneme condition. Interestingly, [ˀt], the variant which Chambers (2009) reports is negatively perceived, was given the lowest rating, although this was not significant. The difference in rating between the two allophones was smaller than the difference between the nasal and approximant contrasts, suggesting that the two were perceived similarly, with similar strength of category-goodness. This supports the finding from Task 2; the [t- ˀt] contrast appears to be the most difficult to perceive.

Therefore, this evidence does not support the hypothesis; despite the presence of [t- ˀt] in the accents of the participants, it was not the most accurately perceived pairing. This may be due to the status of the contrast in the native language. Iverson et al. (2003) researched contrasts with phonemic status and, although Kelly (2017) investigated the perception of allophones, the pairing, dental-alveolar stops, were contrastive for Irish, whereas the contrast used in this study was within-category. So, it may be that, while accent and language experience influences between category perception, it does not influence perception within category.

Suggestions for further research

Methodological evaluation

There were some methodological limitations with the study. Firstly, Task 1 was repetitive and some of the younger participants struggled to understand the instructions at first. To resolve this, practice trials could have been used to test understanding of the instructions before beginning the control condition. Alternatively, a different method, like the one used in Ohde and Haley (1997) could have been used which, as well as being more engaging, may be easier for the child participants to understand.

Likewise, if the experiment was repeated, instead of using the letter chart for Task 2, participants could be asked to write the letter they identify the stimuli as either on a response sheet or whiteboard. This is because some of the participants were confused by the font used in the chart as the descender on <g> was different to the style they were used to and as, even when the participants were able to identify the stimuli, some struggled to find the corresponding box on the chart. Using a whiteboard would overcome these issues, although it may increase the variety in responses.

The study found that form was significant in several of the analyses. Therefore, to reduce the effect of form, the stimuli should be presented in a random order for each participant, rather than in a pre-set order as this may reduce the impact of order of presentation on the findings.

Further studies

With these modifications, the experiment could be conducted with different participant groups. Firstly, the experiment could be conducted with a different accent group, such as with participants from Leeds or Newcastle as Carter and Local (2007) note that these varieties differ in their distribution of dark and clear /l/. The study could be conducted to test whether the findings for the [l-ɫ] contrast are also found with these groups. To test the effects of accent, a follow up questionnaire could also be used after the testing to evaluation participants’ awareness of the contrasts.

The experiment could also be conducted with participants diagnosed with dyslexia. This is because several studies suggest that children with dyslexia perceive differently to typically developing children (eg. Bogliotti, Serniclaes, Messaoud-Galusi & Sprenger-Charolles, 2008; Manis et al., 1997; Snowling, Lervåg, Nash & Hulme, 2018). One study, by Serniclaes et al. (2004) suggests that while typically developing children perceive phonemically, children with dyslexia show “allophonic perception” which “disrupts the one-to-one relation between letters and sounds” (Serniclaes et al. 2004, p.357).

So, a study based on this project could be used to answer Sernicales et al.’s (2004) question of whether allophonic perception changes with age in children with dyslexia. This could have implications for speech and language therapy as some approaches to therapy include training on phonemic awareness (Zoubrinetzky, Collet, Nguyen-Morel, Valdois & Serniclaes, 2019). So, if allophonic perception is found to occur throughout development, this could also be addressed in therapy, for example, identifying sounds as belonging to the same category.

Conclusion

To conclude, this project researched the development of perception in children between the ages of 6-15-years-old in comparison to an adult control group. To do this, the research used three tasks; an AX discrimination task, an identification task, and a rating task. The study found that both allophone and phoneme perception increased with age. This could be due to the increase in experience with the native language with age which may have a faciliatory effect on perception (based on Iverson et al. 2003; Kelly, 2017). However, this change could also be due to non-linguistic factors such as the development of memory abilities (Swets (1964, cited in Sussman & Carney, 1989). The study found that development did not occur in stages, although there appear to be changes in perception around the age of 10-11-years-old, which could be linked to the critical period, although there are issues with this explanation.

As well as considering age, this research investigated the proposal that gestures could be used as the unit of perception, as found in Liberman and Mattingly's (1985) Motor Theory and Best and McRobert's (2003) Perceptual Assimilation Model. The findings support this proposal; the greater the difference in articulatory gesture, the stronger the perception abilities as [l-ɫ] was more accurately discriminated than [n-n̪]. This supports the predictions made by PAM; [l-ɫ] may have been easier to discriminate as the contrast may have been perceived as a two-category contrast, as participants made more errors in the identification task, or a category-goodness contrast, as [ɫ] was rated lower than the other allophones. On the other hand, the nasal allophone contrast may have been more difficult to discriminate as the contrast may have been perceived as a single-category contrast, which, according to PAM, are more difficult to discriminate, as [n] and [n̪] were identified and rated similarly.

This project also considered the influence of accent on perception, as the literature suggests that language experience influences perception abilities (Iverson et al., 2003; Kelly, 2017). This research did not find an effect of accent in perception. This may be because the pairing used in this study, [t-ˀt], was not contrastive and so, despite being a feature of the accent of the participants, was difficult to discriminate.

Finally, this research could be repeated, with modifications, with different participant groups such as different accent groups to investigate the influence of accent further, or with children diagnosed with dyslexia to investigate the proposal that children with dyslexia show allophonic rather than phonemic perception.

References

Altmann, G. (1998). *The Ascent of Babel*. Oxford: Oxford University Press

Best, C. & McRoberts, G. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech*, *46*(2-3), 183-216.

Best, C. (1991). The emergence of native-language phonological influences in infants: a perceptual assimilation model. Haskins laboratories Status Report of Speech Research, SR-107-108, 1-30.

Bogliotti, C., Serniclaes, W., Messaoud-Galusi, S., & Sprenger-Charolles, L. (2008). Discrimination of speech sounds by children with dyslexia: Comparisons with chronological age and reading level controls. *Journal of Experimental Child Psychology, 101*(2), 137-155.

Carter, P., & Local, J. (2007). F2 variation in Newcastle and Leeds English liquid systems. Journal of the International Phonetic Association, 37(2), 183-199.

Chambers, J. (2009). *Sociolinguistic theory* (revised ed.). Chichester: Wiley-Blackwell Kelly.

Choi, D., Bruderer, A. G., & Werker, J. F. (2019). Sensorimotor influences on speech perception in pre-babbling infants: Replication and extension of Bruderer et al. (2015). *Psychonomic Bulletin & Review, 26*(4), 1388-1399.

D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009). The motor somatotopy of speech perception. *Current Biology, 19*(5), 381-385.

Eimas, P. (1974). Auditory and linguistic processing of cues for place of articulation by infants. *Perception & Psychophysics, 16*(3), 513-521.

Eimas, P. (1975). Auditory and phonetic coding of the cues for speech: Discrimination of the [r-l] distinction by young infants. *Perception & Psychophysics, 18*(5), 341-347.

Fikkert, P. (1994). *On the acquisition of prosodic structure*. The Hague: Holland Institute of Generative Linguistics. Retrieved from https://repository.ubn.ru.nl//bitstream/handle/2066/32125/on__thaco.pdf

Galantucci, B., Fowler, C., & Turvey, M. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review, 13*(3), 361-377.

Harnsberger, J. D. (2001). On the relationship between identification and discrimination of non-native nasal consonants. *The Journal of the Acoustical Society of America, 110*(1), 489-503.

Hayes, B. (2009). *Introductory phonology*. Chichester: Blackwell Publishing.

Hohne, E., & Jusczyk, P. (1994). Two-month-old infants' sensitivity to allophonic differences. *Perception & Psychophysics, 56*(6), 613-623.

Iverson, P., Kuhl, K., Akahana-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. & Siebter, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*(1), pp.B47-B57.

Johnson, K., & Babel, M. (2010). On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of Phonetics, 38*(1), 127-136.

Jolly Phonics. (n.d.). Retrieved from https://www.jollylearning.co.uk/jolly-phonics/

Jusczyk, P. W., & Thompson, E. (1978). Perception of a phonetic contrast in multisyllabic utterances by 2-month-old infants. *Perception & Psychophysics, 23*(2), 105-109.

Jusczyk, P. W., Copan, H., & Thompson, E. (1978). Perception by 2-month-old infants of glide contrasts in multisyllabic utterances. *Perception & Psychophysics, 24*(6), 515-520.

Jusczyk, P., Shea, S. & Aslin, E. (1984). Linguistic experience and infant speech perception: a re-examination of Eilers, Gavin & Oller (1982). *Journal of Child Language, 11*(2), 453-466

Kelly, N. (2017). The perception of dental and alveolar stops among speakers of Irish English and American English. *English Language and Linguistics, 23*(2), 277-302.

Larkey, L., Wald, J., & Strange, W. (1978). Perception of synthetic nasal consonants in initial and final syllable position. *Perception & Psychophysics, 23*(4), 299-312.

Liberman, A., & Mattingly, I. (1985). The motor theory of speech perception revised. *Cognition, 21*(1), 1-36.

Locke, J. (1980). The inference of speech perception in the phonologically disordered child. Part II: Some clinically novel procedures, their use, some findings. *Journal of Speech and Hearing Disorders, 45*(4), 445-468.

Manis, F., McBride-Chang, C., Seidenberg, M., Keating, P., Doi, L., Munson, B., & Petersen, A. (1997). Are speech perception deficits associated with developmental dyslexia? *Journal of Experimental Child Psychology, 66*(2), 211-235.

Moffitt, A. (1971). Consonant cue perception by twenty-to twenty-four-week-old infants. *Child Development, 42*(3) 717-731.

Narayan, C., Werker, J., & Beddor, P. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science, 13*(3), 407-420.

Ohde, R. & Haley, K. (1997). Stop-consonant and vowel perception in 3- and 4-year-old children. *The Journal of the Acoustical Society of America, 102*(6), 3711-3722.

Peperkamp, S., Pettinato, M., & Dupoux, E. (2003). Allophonic variation and the acquisition of phoneme categories. *Proceedings of the 27th Annual Boston University Conference on Language Development, 2*, 650–661. Sommerville, MA : Cascadilla Press.

Ruben, R. (1997). A time frame of critical/sensitive periods of language development. *Acta oto-laryngologica*, *117*(2), 202-205.

Seidl, A., Cristià, A., Bernard, A., & Onishi, K. (2009). Allophonic and phonemic contrasts in infants' learning of sound patterns. *Language Learning and Development*, *5*(3), 191-202.

Serniclaes, W., Van Heghe, S., Mousty, P., Carré, R., & Sprenger-Charolles, L. (2004). Allophonic mode of speech perception in dyslexia. *Journal of Experimental Child Psychology*, *87*(4), 336-361.

Singleton, D. (2005). The critical period hypothesis: A coat of many colours. *International Review of Applied Linguistics in Language Teaching*, *43*(4), 269-285.

Snow, C., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child development*, 49(4)1114-1128.

Snowling, M., Lervåg, A., Nash, H., & Hulme, C. (2019). Longitudinal relationships between speech perception, phonological skills and reading in children at high-risk of dyslexia. *Developmental Science*, *22*(1), 1-12.

Sussman, J. & Carney, A. (1989). Effects of transition length on the perception of stop consonants by children and adults. *Journal of Speech and Hearing Research, 32*(1), 151-160.

Trudgill, P. (1974). *The social differentiation of English in Norwich*. London: Cambridge University Press.

Whalen, D., Best, C., & Irwin, J. (1997). Lexical effects in the perception and production of American English/p/allophones. *Journal of Phonetics, 25*(4), 501-528.

Zoubrinetzky, R., Collet, G., Nguyen-Morel, M., Valdois, S. & Serniclaes, W. (2019). Remediation of allophonic perception and visual attention span in developmental dyslexia: a joint assay. *Frontiers in Psychology, 10*(1502), 1-16.

Zsiga, E. (2013). *The sounds of language: an introduction to phonetics and phonology.* Chichester: Blackwell Publishing.

Appendices

Appendix A: Response form instructions

1. Are you happy to take part in the experiment today?
2. You will be given three tasks. I will explain each task before it starts.
3. I will offer a break at least every 15 trials. If you feel you want to take a break at any time, please let me know.
   We can play a game of Jenga while we do the experiment. Did you want to play a game?
4. I will play the sounds over these headphones which have been cleaned. Please make sure they're comfortable before we begin.
5. I will play a test sound now. Is this loud enough?
6. Are you ready to begin?

Task 1:

1. I am going to play you pairs of sounds.
2. Each sound is made of a vowel, consonant vowel pattern like /apa/-/aka/. Do you know what vowels and consonants are?
3. Thinking about the consonants, please say whether the second sound is the same or different to the first.

Code: *Same: A* *Different: B*

Task 2:

1. Can you read all of the sounds on this chart?
2. I am going to play you more sounds.
3. When I play the sound, please point to the box on the chart the sound is.

Note the graphemes the participant identifies the stimuli as.

Task 3:

1. I am going to play you a sound more sounds.
2. Please rate how normal the sound is on a scale.
3. For example, if I said make a /p/ sound, is the /p/ and normal one or slightly strange.
4. If the sound is a typical example, it would be given a rating of 5. If the sound is an abnormal example of the sound, it would be given a rating of 1.
   If the sound sounds like a normal sound you would say, please point to the smiley face. If the sound sounds a bit weird it would get a straight face.
5. When I play the sound, please point to the symbol.

Code: 1 to 5 for participants aged above 7.

For participants below 7-years of age:

B- atypical (bad) production *G- typical (good) production*

Appendix B: Table 9: Participant response form

Participant ID		**Gender**	
Year		**Form**	

Task 1								**Task 2**				**Task 3**			
1		31		1		31		1		1		1		1	
2		32		2		32		2		2		2		2	
3		33		3		33		3		3		3		3	
4		34		4		34		4		4		4		4	
5		35		5		35		5		5		5		5	
6		36		6		36		6		6		6		6	
7		37		7		37		7		7		7		7	
8		38		8		38		8		8		8		8	
9		39		9		39		9		9		9		9	
10		40		10		40		10		10		10		10	
11		41		11		41		11		11		11		11	
12		42		12		42		12		12		12		12	
13		43		13		43		13		13		13		13	
14		44		14		44		14		14		14		14	
15		45		15		45		15		15		15		15	
16		46		16		46		16		16		16		16	
17		47		17		47		17		17		17		17	
18		48		18		48		18		18		18		18	
19		49		19		49									
20		50		20		50									
21		51		21		51									
22		52		22		52									
23		53		23		53									
24		54		24		54									
25		55		25		55									
26		56		26		56									
27		57		27		57									
28		58		28		58									
29		59		29		59									
30		60		30		60									

Participant ID		Gender	
Year		Form	

TASK 1								TOTAL
	Trial	Stimuli	1	2	3	4	5	**Correct**
CONTROL	1	[m-m]						
	2	[ŋ-ŋ]						
	3	[m-ŋ]						
	4	[ŋ-m]						
	5	[d-d]						
	6	[g-g]						
	7	[d-g]						
	8	[g-d]						
	9	[j-j]						
	10	[ɹ- ɹ]						
	11	[j- ɹ]						
	12	[ɹ-j]						
EXPERIMENTAL	1	[n-n]						
	2	[n̪-n̪]						
	3	[n-n̪]						
	4	[n̪-n]						
	5	[t- t]						
	6	[ˀt-ˀt]						
	7	[t-ˀt]						
	8	[ˀt- t]						
	9	[l-l]						
	10	[ɫ-ɫ]						
	11	[l-ɫ]						
	12	[ɫ-l]						

TASK 2						Modal score
	Trial	Stimuli	1	2	3	
CONTROL	1	/m/				
	2	/ŋ/				
	3	/d/				
	4	/g/				
	5	/j/				
	6	/ɹ/				
EXPERIMENTAL	9	[n]				
	10	[n̪]				
	11	[t]				
	12	[ˀt]				
	13	[l]				
	14	[ɫ]				

	TASK 3					Average score
	Trial	Stimuli	1	2	3	
CONTROL	1	/m/				
	2	/ŋ/				
	3	/d/				
	4	/g/				
	5	/j/				
	6	/ɹ/				
EXPERIMENTAL	9	[n]				
	10	[n̪]				
	11	[t]				
	12	[ˀt]				
	13	[l]				
	14	[ɫ]				

Appendix C: Identification chart

Based on the Jolly Phonics system (Jolly Phonics, n.d.)

atha	**ata**	**aba**	**ana**	**awa**	**acha**

ara	**ama**	**apa**	**aga**	**acka**	**aya**

aja	**afa**	**ala**	**aha**	**ada**	**anga**

Figure 1: Identification chart

Appendix D: SPSS output tables for Analysis 1A

Binomial regression on Task 1 data considering participant category, gender, form and stimuli code.

Model Summary	
-2 Log likelihood	Nagelkerke R Square
4633.063[a]	.690

Table 10: SPSS Model Summary for Analysis 1A

Classification Table				
		Predicted		
		Accuracy		
Observed		0	1	Percentage Correct
Accuracy	0	1568	643	70.9
	1	152	7957	98.1
Overall Percentage				92.3

Table 11: SPSS Classification Table for Analysis 1A

Variables in the Equation						
	B	S.E.	Wald	df	Sig.	Exp(B)
Participant category	-.490	.131	13.886	1	.000	.613
[m-m]			2817.442	23	.000	
[m-ŋ]	-.926	.402	5.311	1	.021	.396
[ŋ-m]	-.463	.297	2.427	1	.119	.629
[ŋ-ŋ]	.738	.388	3.611	1	.057	2.091
[d-d]	-.160	.293	.296	1	.586	.852
[d-g]	-1.286	.200	41.274	1	.000	.276
[g-d]	-1.208	.183	43.633	1	.000	.299
[g-g]	.852	.350	5.914	1	.015	2.344
[j-j]	.450	.308	2.144	1	.143	1.569
[j- ɹ]	-.323	.228	1.999	1	.157	.724
[ɹ-j]	-.291	.226	1.653	1	.199	.748
[ɹ- ɹ]	.228	.274	.693	1	.405	1.257
[n-n]	.627	.329	3.641	1	.056	1.872
[n-n̪]	-5.862	.210	782.512	1	.000	.003
[n̪-n]	-4.760	.164	839.812	1	.000	.009
[n̪-n̪]	1.679	.387	18.845	1	.000	5.361
[t- t]	1.112	.313	12.660	1	.000	3.041
[t-ˀt]	-5.024	.186	725.815	1	.000	.007
[ˀt- t]	-4.745	.185	654.468	1	.000	.009
[ˀt-ˀt]	1.781	.342	27.101	1	.000	5.938

[l-l]	2.105	.415	25.680	1	.000	8.208
[l-ɫ]	-1.902	.115	273.220	1	.000	.149
[ɫ-l]	-1.555	.116	178.466	1	.000	.211
[ɫ-ɫ]	1.369	.287	22.733	1	.000	3.930
Gender	-.121	.080	2.281	1	.131	.886
Form A			9.572	4	.048	
Form B	.083	.115	.523	1	.470	1.087
Form C	.160	.106	2.267	1	.132	1.173
Form D	-.267	.104	6.548	1	.010	.766
Form E	.090	.116	.597	1	.440	1.094
Constant	2.668	.137	379.782	1	.000	14.410

Table 12: SPSS Variables in the Equation for Analysis 1A

Appendix E: SPSS output tables for Analysis 1B

Binomial regression on child data from Task 1 considering age, gender, form and stimuli code.

Model Summary	
-2 Log likelihood	Nagelkerke R Square
3950.233[a]	.707

Table 13: SPSS Model Summary for Analysis 1B

Classification Table				
Observed		Predicted		
		Accuracy		Percentage Correct
		0	1	
Accuracy	0	1477	507	74.4
	1	193	6943	97.3
Overall Percentage				92.3

Table 14: SPSS Classification Table for Analysis 1B

Variables in the Equation						
	B	S.E.	Wald	df	Sig.	Exp(B)
[m-m]			2485.760	23	.000	
[m-ŋ]	-.884	.406	4.740	1	.029	.413
[ŋ-m]	-.336	.311	1.163	1	.281	.715
[ŋ-ŋ]	.675	.391	2.984	1	.084	1.963
[d-d]	-.211	.296	.512	1	.474	.809
[d-g]	-1.161	.212	30.078	1	.000	.313
[g-d]	-1.023	.198	26.576	1	.000	.360
[g-g]	.767	.352	4.757	1	.029	2.154
[j-j]	.374	.309	1.469	1	.226	1.454
[j- ɹ]	.033	.272	.015	1	.903	1.034
[ɹ-j]	.169	.286	.350	1	.554	1.185
[ɹ- ɹ]	.611	.348	3.089	1	.079	1.842
[n-n]	.817	.390	4.387	1	.036	2.264
[n-n̪]	-6.062	.231	686.922	1	.000	.002
[n̪-n]	-5.080	.189	723.033	1	.000	.006
[n̪-n̪]	1.526	.389	15.429	1	.000	4.601
[t- t]	1.173	.345	11.554	1	.001	3.233
[t-ˀt]	-5.168	.204	638.958	1	.000	.006
[ˀt- t]	-4.918	.206	569.357	1	.000	.007
[ˀt-ˀt]	1.650	.344	23.033	1	.000	5.209

[l-l]	1.982	.417	22.626	1	.000	7.260
[l-ɫ]	-2.093	.122	292.526	1	.000	.123
[ɫ-l]	-1.731	.123	198.717	1	.000	.177
[ɫ-ɫ]	1.264	.289	19.166	1	.000	3.540
Gender	-.184	.087	4.492	1	.034	.832
Form A			7.382	4	.117	
Form B	.191	.124	2.377	1	.123	1.211
Form C	.024	.114	.043	1	.835	1.024
Form D	-.251	.114	4.856	1	.028	.778
Form E	-.041	.131	.099	1	.753	.960
Age	.114	.017	46.840	1	.000	1.121
Constant	1.008	.190	28.041	1	.000	2.741

Table 15: SPSS Variables in the Equation for Analysis 1B

Appendix F: SPSS output tables for Analysis 1C

Binomial regression on Task 1, considering participant category, stimuli category, gender and form.

Model Summary	
-2 Log likelihood	Nagelkerke R Square
8897.331[a]	.251

Table 16: SPSS Model Summary for Analysis 1C

Classification Table				
		Predicted		
		Accuracy		
Observed		0	1	Percentage Correct
Accuracy	0	0	2211	.0
	1	0	8109	100.0
Overall Percentage				78.6

Table 17: SPSS Classification Table for Analysis 1C

Variables in the Equation						
	B	S.E.	Wald	df	Sig.	Exp(B)
Participant category	-.207	.084	6.035	1	.014	.813
Gender	-.052	.053	.992	1	.319	.949
Form A			4.169	4	.384	
Form B	.036	.076	.227	1	.634	1.037
Form C	.069	.069	.978	1	.323	1.071
Form D	-.117	.069	2.839	1	.092	.889
Form E	.039	.076	.267	1	.606	1.040
Stimuli category	-2.436	.070	1220.087	1	.000	.087
Constant	1.927	.087	486.955	1	.000	6.866

Table 18: SPSS Variables in the Equation for Analysis 1C

Appendix G: SPSS output tables for Analysis 2A

Binomial regression on Task 2 data considering participant category, gender, form and stimuli code.

Model Summary	
-2 Log likelihood	Nagelkerke R Square
1808.477[a]	.128

Table 19: SPSS Model Summary for Analysis 2A

Classification Table				
Observed		Predicted		
		Accuracy		Percentage Correct
		0	1	
Accuracy	0	0	306	.0
	1	0	2826	100.0
Overall Percentage				90.2

Table 20: SPSS Classification Table for Analysis 2A

Variables in the Equation						
	B	S.E.	Wald	df	Sig.	Exp(B)
/m/			152.253	11	.000	
/ŋ/	-1.799	.423	18.094	1	.000	.165
/d/	-.899	.277	10.580	1	.001	.407
/g/	.269	.294	.836	1	.361	1.309
/j/	-.635	.223	8.123	1	.004	.530
/ɹ/	.614	.306	4.037	1	.045	1.849
[n]	-.544	.212	6.590	1	.010	.581
[n̪]	.128	.251	.261	1	.609	1.137
[t]	1.080	.370	8.542	1	.003	2.946
[ˀt]	.168	.266	.399	1	.528	1.183
[l]	1.654	.511	10.482	1	.001	5.226
[ɫ]	-1.789	.163	120.820	1	.000	.167
Gender	-.062	.125	.241	1	.623	.940
Form A			15.024	4	.005	
Form B	-.230	.182	1.599	1	.206	.795
Form C	.239	.177	1.835	1	.176	1.270
Form D	-.516	.153	11.396	1	.001	.597
Form E	.162	.188	.738	1	.390	1.176

Participant category	.270	.183	2.183	1	.140	1.310
Constant	2.328	.188	153.271	1	.000	10.258

Table 21: SPSS Variables in the Equation for Analysis 2A

Appendix H: SPSS output tables for Analysis 2B

Binomial regression on child data from Task 2 considering age, gender, form and stimuli code.

Model Summary	
-2 Log likelihood	Nagelkerke R Square
1563.509[a]	.132

Table 22: SPSS Model Summary for Analysis 2B

Classification Table				
Observed		Predicted		
		Accuracy		Percentage Correct
		0	1	
Accuracy	0	0	263	.0
	1	0	2509	100.0
Overall Percentage				90.5

Table 23: SPSS Classification Table for Analysis 2B

Variables in the Equation						
	B	S.E.	Wald	df	Sig.	Exp(B)
/m/			116.030	11	.000	
/ŋ/	-1.766	.427	17.076	1	.000	.171
/d/	-.813	.286	8.067	1	.005	.444
/g/	.423	.318	1.770	1	.183	1.527
/j/	-.752	.229	10.837	1	.001	.471
/ɹ/	.547	.309	3.136	1	.077	1.729
[n]	-.395	.230	2.948	1	.086	.674
[n̪]	.513	.303	2.868	1	.090	1.670
[t]	.961	.372	6.657	1	.010	2.614
[ˀt]	.188	.283	.439	1	.508	1.207
[l]	1.550	.513	9.144	1	.002	4.711
[ɫ]	-1.558	.180	74.876	1	.000	.211
Gender	-.007	.135	.002	1	.961	.993
Form A			16.676	4	.002	
Form B	-.391	.189	4.277	1	.039	.676
Form C	.434	.196	4.884	1	.027	1.544
Form D	-.500	.168	8.872	1	.003	.607
Form E	.016	.217	.005	1	.943	1.016

Age	.122	.025	23.301	1	.000	1.129
Constant	1.253	.284	19.456	1	.000	3.501

Table 24: SPSS Variables in the Equation for Analysis 2B

Appendix I: SPSS output tables for Analysis 2C

Binomial regression on Task 2, considering participant category, stimuli category, gender and form.

Model Summary	
-2 Log likelihood	Nagelkerke R Square
1987.875[a]	.011

Table 25: SPSS Model Summary for Analysis 2C

Classification Table				
		Predicted		
		Accuracy		
Observed		0	1	Percentage Correct
Accuracy	0	0	306	.0
	1	0	2826	100.0
Overall Percentage				90.2

Table 26: SPSS Classification Table for Analysis 2C

Variables in the Equation						
	B	S.E.	Wald	df	Sig.	Exp(B)
Gender	-.058	.121	.226	1	.635	.944
Form A			14.099	4	.007	
Form B	-.216	.176	1.501	1	.220	.806
Form C	.226	.172	1.724	1	.189	1.253
Form D	-.481	.147	10.702	1	.001	.618
Form E	.150	.183	.678	1	.410	1.162
Participant Category	.252	.176	2.044	1	.153	1.286
Stimuli category	-.087	.121	.524	1	.469	.916
Constant	2.047	.175	137.398	1	.000	7.744

Table 27: SPSS Variables in the Equation for Analysis 2C

Appendix J: SPSS output tables for Analysis 3A

Generalized linear model on Task 3 data considering participant category, gender, form and stimuli code.

Omnibus Test		
Likelihood Ratio Chi-Square	df	Sig.
252.105	17	.000

Table 28: SPSS Omnibus Test for Analysis 3A

Parameter Estimates								
Parameter		B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
				Lower	Upper	Wald Chi-Square	df	Sig.
Threshold	[Response=1]	.072	.0888	-.102	.246	.652	1	.419
	[Response=2]	.329	.0889	.154	.503	13.666	1	.000
	[Response=3]	.696	.0896	.520	.871	60.295	1	.000
	[Response=4]	1.171	.0907	.993	1.349	166.772	1	.000
/m/		.879	.0984	.686	1.072	79.786	1	.000
/ŋ/		.386	.0970	.196	.576	15.830	1	.000
/d/		1.018	.0999	.822	1.214	103.860	1	.000
/g/		.763	.0980	.571	.955	60.589	1	.000
/j/		.558	.0972	.367	.748	32.912	1	.000
/ɹ/		.766	.0982	.573	.958	60.838	1	.000
[n]		.984	.0988	.790	1.178	99.174	1	.000
[n̪]		.798	.0981	.606	.990	66.203	1	.000
[t]		.751	.0980	.558	.943	58.617	1	.000
[ˀt]		.644	.0977	.452	.835	43.381	1	.000
[l]		.512	.0976	.321	.703	27.531	1	.000
[ɫ]		0[a]	.	.	.	.	.	.

Gender (Female)	.114	.0398	.036	.192	8.147	1	.004
Gender (Male)	0[a]	.	.	.	.	.	.
Form A	.112	.0671	-.020	.243	2.763	1	.096
Form B	.056	.0665	-.075	.186	.697	1	.404
Form C	-.212	.0684	-.346	-.078	9.597	1	.002
Form D	-.032	.0705	-.170	.107	.200	1	.654
Form E	0[a]	.	.	.	.	.	.
Child participants	.366	.0619	.244	.487	34.911	1	.000
Adult Participants	0[a]	.	.	.	.	.	.

Table 29: SPSS Parameter Estimates for Analysis 3A

Appendix K: SPSS output tables for Analysis 3B

Generalised linear model on child data from Task 3 considering age, gender, form and stimuli code.

Omnibus Test		
Likelihood Ratio Chi-Square	df	Sig.
252.570	17	.000

Table 30: SPSS Omnibus Test for Analysis 3B

Parameter Estimate								
Parameter		B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
				Lower	Upper	Wald Chi-Square	df	Sig.
Threshold	[Response=1]	.972	.1389	.700	1.244	48.949	1	.000
	[Response=2]	1.210	.1403	.935	1.485	74.373	1	.000
	[Response=3]	1.569	.1418	1.291	1.847	122.491	1	.000
	[Response=4]	2.021	.1419	1.743	2.299	202.949	1	.000
/m/		.813	.1050	.607	1.019	59.955	1	.000
/ŋ/		.315	.1042	.110	.519	9.114	1	.003
/d/		.922	.1064	.713	1.131	75.058	1	.000
/g/		.695	.1047	.490	.901	44.080	1	.000
/j/		.484	.1041	.280	.688	21.652	1	.000
/ɹ/		.664	.1049	.459	.870	40.106	1	.000
[n]		.908	.1054	.702	1.115	74.244	1	.000
[n̪]		.701	.1047	.495	.906	44.748	1	.000
[t]		.623	.1046	.417	.828	35.415	1	.000
[ˀt]		.529	.1043	.324	.733	25.716	1	.000

[l]	.387	.1044	.183	.592	13.774	1	.000
[ł]	0[a]	.	.	.	.	.	.
Gender (Female)	.070	.0426	-.014	.153	2.678	1	.102
Gender (Male)	0[a]	.	.	.	.	.	.
Form A	.300	.0731	.156	.443	16.819	1	.000
Form B	.209	.0721	.068	.351	8.436	1	.004
Form C	-.066	.0744	-.212	.080	.781	1	.377
Form D	.207	.0773	.056	.359	7.191	1	.007
Form E	0[a]	.	.	.	.	.	.
Age	.074	.0084	.057	.090	77.333	1	.000

Table 31: SPSS Parameter Estimates for Analysis 3B

Appendix L: SPSS output tables for Analysis 3C

Generalized linear model on Task 3, considering participant category, stimuli category, gender and form.

Omnibus Test		
Likelihood Ratio Chi-Square	df	Sig.
79.489	7	.000

Table 32: SPSS Omnibus Test for Analysis 3C

Parameter Estimates								
Parameter		B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
				Lower	Upper	Wald Chi-Square	df	Sig.
Threshold	[Response=1]	-.522	.0623	-.644	-.400	70.185	1	.000
	[Response=2]	-.274	.0618	-.395	-.153	19.625	1	.000
	[Response=3]	.076	.0617	-.045	.197	1.498	1	.221
	[Response=4]	.532	.0624	.410	.655	72.898	1	.000
Gender (Female)		.111	.0396	.033	.189	7.820	1	.005
Gender (Male)		0[a]	.	.	.	.	.	.
Form A		.110	.0668	-.021	.241	2.716	1	.099
Form B		.052	.0662	-.078	.182	.612	1	.434
Form C		-.203	.0681	-.337	-.070	8.902	1	.003
Form D		-.032	.0701	-.169	.106	.203	1	.652
Form E		0[a]	.	.	.	.	.	.
Child participants		.351	.0614	.231	.472	32.716	1	.000
Adult Participants		0[a]	.	.	.	.	.	.
Phoneme condition		.107	.0395	.029	.184	7.273	1	.007
Allophone condition		0[a]	.	.	.	.	.	.

Table 33: SPSS Parameter Estimates for Analysis 3C