Audiovisual Speech Perception in Children with Normal and Late Language Emergence

Rafiah Badat

City University London

Target Journal: *Journal of Speech, Language and Hearing Research*

Dissertation submitted in partial fulfilment of the MSc in Human Communication
Department of Language and Communication Science, City University, London

June, 2014

Declaration:

I grant powers of discretion to the Department of Language and Communication Science to allow this dissertation to be copied in whole or in part without any further reference to me. This permission covers only single copies made for study purposes, subject to the normal conditions of acknowledgement.

Author's Note

This research was supported by Dr. Victoria Knowland, Research Fellow, Division of Language and Communication Science, School of Health Science, City University London

Correspondence concerning this article should be addressed to:
rafiah.badat.1@city.ac.uk

Abstract

Purpose: To investigate the audiovisual perception of novel phonological information and the influence of processing demands in children with late language emergence.

Method: Fifty primary-school-aged children, half with typically developing language and half with delayed language acquisition, were asked to repeat non-words varying in syllable length and articulation complexity. Presentation alternated randomly between hearing a recording of a speaker and both hearing and seeing a pre-recording.

Results: The late-talking group showed consistently lower repetition accuracy than the children with normal language onset regardless of non-word presentation; however differences generally failed to reach significance. The late-talking children benefitted from visual cuing to a significantly greater extent than the typical language emergence group, this was particularly evident with stimuli of lower length and articulation demand.

Conclusions: School-aged children with apparently resolved language needs found repetition of novel word forms more difficult than children without delayed language onset. Visual cues supported verbal perception and processing for late-talking children considerably more than for typical peers, perhaps indicating a more immature language learning mechanism. However, as verbal processing and production demands increased, the influence of visual support for late-talkers decreased reflecting a narrow margin of benefit.

Keywords: late language emergence, late talkers, audiovisual speech perception, non-word repetition

Audiovisual Speech Perception in Children with Normal and Late Language Emergence

The emergence of a child's first language is a complex and individually unique process governed by the dynamic interaction of multiple subsystems within the child, the social context and the physical environment (Darrah, Hodge, Magill-Evans, & Kembhavi, 2003). There are several factors highlighted through large scale epidemiology studies as having a collective influence on both rate and pattern of early language development. The most commonly identified are gender, general health and development, family history and socioeconomic status; however none of the areas studied have shown sufficient individual impact to accurately predict language outcome (Henrichs et al., 2011; Reilly et al., 2010).

Despite the considerable heterogeneity in the trajectory of first language acquisition, the broad benchmarks of words appearing around the first birthday and use of word combinations by the second birthday are widely accepted (Darley & Winitz, 1961; Preston et al., 2010). Empirical studies indicate that in around 15% of preschool-aged children these early language milestones are delayed despite the absence of an identifiable deficit in the areas of sensory, cognitive, physical or social development (Broomfield & Dodd, 2004; McLeod & Harrison, 2009; Reilly et al., 2009). Further investigations show that by school-age up to three quarters of the children with late language emergence (LLE) present within the normal rage on standardised assessments of vocabulary and grammar (Paul and Roth, 2011). Despite the progress made, these LLE children cannot be considered to have truly resolved needs given that they tend to consistently fall in the low average bracket on language norms and often have marked difficulties with conversational language (Domsch et al., 2012; Rescorla, 2011). Furthermore, corresponding weaknesses have been found to persist well into adolescence with literacy, academic attainment, psychosocial wellbeing and functional independence (Preston et al., 2010; Shevell, Majnemer, Webster, Platt & Birnbaum, 2005; Snowling, Adams, Bishop & Stothard 2001).

Language Impairment and NWR

There currently exists a substantial body of research focused on identifying the mechanisms underlying paediatric language impairment. The vast majority of these studies have centred on children with isolated language difficulties who fail to progress to the statistically defined normal language range, a phenomenon commonly referred to as a specific language impairment (SLI; Tomblin, Records & Zhang, 1996). Children with SLI form an extremely diverse group and difficulties may manifest in one or more areas of receptive and expressive language including phonology, vocabulary, morpho-syntax, sentence structure and pragmatics (Marinis, 2011). Unsurprisingly, concomitant difficulties with learning and quality of life are more prominent than in those late-talking children who do not go on to present with the features of SLI, perhaps indicating a continuum of severity (Rescorla, 2002).

One particularly robust research finding is the extent to which weak non-word repetition (NWR) skills characterises not only the SLI population (Estes, Evans, & Else-Quest, 2007) but also children with LLE who go on to develop language within normal limits (Bishop, North & Donlan, 1996; Conti-Ramsden, Botting & Faragher, 2001). This suggests that NWR deficits may be a broad phenotype for weak language learning capacity with language outcomes varying depending on the nature and extent of the underlying difficulty.

The act of repeating a non-word is a complex and multifaceted process involving the perception, short term storage, processing and articulation of novel phonological forms (Bishop, 2006). It is thought that children with language difficulties may experience deficits either at a particular point or at multiple points within this NWR sequence (Coady & Evans, 2008).

There are numerous theories that attempt to explain how poor NWR skills relate to restricted language progression. One of the earliest is based on the fact that both SLI and

LLE populations demonstrate markedly greater difficulty with NWR of structures beyond 2 syllables, which contrasts with a more gradual decline in performance with typical-language control groups (Bishop et al., 1996; Gathercole & Baddelay, 1990). It is proposed that this demonstrates weak short-term phonological memory amongst individuals with language learning deficits. The result is a reduced capacity to temporarily store and rehearse new word forms for embedding into long-term lexical memory, thereby leading to a restricted vocabulary (Gathercole, 2006). This constraint on the range and accuracy of phonological representations caused by a limited phonological memory may also impact on grammatical development due to poor retention of morphological inflections (Marinis, 2011).

Manipulation of non-words outside the syllable level provides further insight into the link between NWR and linguistic skills. For example, both consonant clusters and developmentally later acquired sounds can lead to a greater reduction in NWR performance in LLE and SLI children than in typically developing peers even in the absence of overt speech production errors (Bishop et al., 1996; Estes et al., 2007). As there is not a direct impact on memory load, this has led to the suggestion that, in addition to a short-term memory deficit, language learning difficulties may also be characterised by inefficiencies with phoneme perception and production. (Estes et al., 2007; Moore, Tompkins & Dollaghan, 2010).

The phonotactic probability, that is the frequency with which a phonological sequence occurs in a given language (Jusczyk & Luce, 1994), has also been shown to differentiate normal and disordered language populations with the latter group showing fewer benefits from linguistic familiarity during NWR tasks (Boelhouwer, 2011; MacRoy-Higgins, Schwartz, Shafer & Marton, 2013).

According to Chiat (2001) language acquisition involves mapping form to meaning by relating phonological information to the linguistic context. Children with language impairment may therefore struggle not only with phonological retention, planning and production but also with the use of prior lexical knowledge to support the manipulation of new phonological forms. Chiat (2001) argues that, in addition to a reduced linguistic range in terms of vocabulary and grammatical morphemes, underlying phonological difficulties may also impact on sentence syntax as word order is dependent on the subtle interactions between content and function words which in turn are defined by their unique phonology.

Given the vast spectrum of presentation within the paediatric language impaired population, it is possible that the disparity in persistency, severity and profile of deficits are due to differences in the source and nature of the phonological breakdown. As such, the multiple theories around NWR and linguistic skills can be considered complimentary explanations for the variations in language outcomes not only between but also within children at different points in time (Bishop, 2006).

Relationship Between NWR and Language Over Time

The trajectory of language development and how it relates to NWR can be considered to be highly informative. Gathercole (2006) has reported strong correlation between NWR and vocabulary skills in typically developing 4-6 year olds with much weaker association by age 8. Similarly, Chiat and Roy (2013) found that although weak NWR at age 2-3 was the strongest predictor of poor morpho-syntax at 4-5 years, by age 10 most of the grammatical deficits had resolved.

These findings led the authors of both studies to propose that first language acquisition is initially highly dependent on a phonological-based learning mechanism, therefore poor phonological manipulation skills as represented by weak NWR are highly correlated with low language levels. However, as lexical knowledge develops, it plays an increasing role in supporting the learning and retaining of further words and syntax patterns with gradually reduced dependency on core phonological skills.

It appears that the persistency of the association between poor NWR skills and low language levels in SLI may be indicative of an ongoing influence of phonological manipulation skills on language learning. In contrast the comparatively stronger lexical competency in LLE may help overcome the impact of underlying phonological deficits. Some evidence of this comes from Gathercole (2006) who makes reference to difficulties with language learning in LLE children that may not be detected on standardised assessments but which can be elicited when semantic and syntactic cues are restricted in order to necessitate reliance on phonological information. This vulnerability to weak language performance due to underlying phonological difficulties amongst the LLE population may explain the risk of low-average outcomes in language, academic attainment and more general functioning in late-talking children who go on to present with superficially adequate linguistic skills.

Audiovisual Integration, NWR and Language Learning

Whilst the majority of studies on isolated language impairment have focused on auditory specific processing of speech stimuli, there are indications that difficulties may extend beyond just the auditory modality and include audiovisual integration.

The ability to correspond auditory and visual information has long been considered to play a role in language acquisition. Infants as young as 10 weeks have been found to pay greater attention to nursery rhymes when what they hear matches what they see in the speaker's face (Dodd, 1979), indicating an early preference for audiovisual synchrony.

At 6 months, the extent to which an infant focuses on their talking mother's mouth has a direct impact on strength of expressive language, vocabulary and social skills at 2 years (Young, Merin, Rogers, & Ozonoff, 2009). By contrast, from 6-9 months it is the preferential focus on the speaker's eyes which predicts language comprehension levels at 14-16 months (Kushnerenko et al., 2013). It is assumed that once the infant has stored phonological information gained by synchronising auditory and visual speech articulation cues, visual priority moves to social information obtained from following the speaker's eyes rather than mouth, a preference that continues into adulthood (Lewkowicz & Hansen-Tift, 2012). This pattern reinforces the theories of Chiat and Roy (2013) and Gathercole (2006) with regards to the developmental progression of language learning from a phonologically focused to a semantically biased mechanism.

It is proposed that after early infancy, visual speech cues provide more compensatory information which come into play when the auditory signal is degraded thus helping to maintain intelligibility (Campbell, 2008). This ability to use visual verbal information to augment auditory cues develops over time with adults showing far greater competency than children (Lewkowicz and Flom, 2013)

It could therefore be argued that vulnerability to weak language acquisition may be influenced by difficulties with audiovisual processing which impacts on the ability to analyse and retain phonological information during infancy and could contribute to persisting difficulties with language development.

Davies, Kidd and Lander (2009) have supported the theory of a link between visual speech perception, phonological processing and early language levels by demonstrating a significant correlation between ability to comprehend a visually presented spoken word, NWR skills and vocabulary size in typically developing children as young as 2. Hosozawa, Tanaka, Shimizu, Nakano and Kitazawa (2012) also investigated the mechanisms of visual verbal perception and found during an eye-tracking study with children aged around 2-3 years that typically developing children tended to shift between a speaker's mouth and eyes, with preferential looking at the mouth only when novel information was introduced. In contrast, children with language difficulties focused predominantly on the speaker's mouth regardless of the familiarity of the speech content. It was suggested that verbal processing difficulties may have restricted the benefit of lexical knowledge for these children leading to excessive focus on the mouth at the expense of gaining social cues from the eyes; however weak audiovisual synchrony may have inhibited the extent to which the additional visual speech cues supported language learning. Furthermore, reduced attention to the speaker's eyes could have led to loss of key social and emotional information resulting in restricted development of social communication skills, as is often characteristic with language disorders (Botting & Conti-Ramsden, 2000; Snowling, Bishop, Stothard, Chipchase, & Kaplan, 2006).

Recent studies have focused more explicitly on the ability to synchronise auditory and visual verbal information in children with a language impairment. Pons, Andreu, Sanz-Torrent, Buil-Legaz, and Lewkowicz (2012) manipulated the audiovisual temporal synchrony of a speaker engaged in conversation and found that Spanish-speaking 4-7 year olds with SLI showed less preferential focus on synchronised than unsynchronised stimuli compared to typically developing controls. It was proposed that the reduced sensitivity to audiovisual disparity may indicate difficulties not only with auditory but also visual speech perception. Similarly, Norrix, Plante, Vance and Boliek (2007) used what is known as the McGurk paradigm to show that a mismatch between auditory and visual presentation influenced accuracy of speech repetition to a significantly greater extent in typically developing preschoolers compared to children with identified language difficulties. Meronen, Tiippana, Westerholm, and Ahonen (2013) found using incongruent audiovisual stimuli that, unlike normal language peers, 6-10 year old SLI children continued to base their NWR responses on what they heard even when the quality of the sound signal was degraded, it was concluded that this led to unsynchronised and therefore less efficient perception and subsequent processing of speech.

The purpose of the current study was to extend findings on audiovisual speech perception by relating to phonological manipulation skills in children with delayed language onset. This was achieved by comparing NWR performance during audio only and audiovisual presentation with the latter involving the speaker being seen as well as heard. It was anticipated that school-aged children with a history of LLE would show a weak language learning capacity as demonstrated by lower NWR scores regardless of whether presentation was audio or audiovisual in comparison to controls with typical language onset. Furthermore, as audiovisual processing difficulties may be a contributor to language learning vulnerabilities in LLE, this group would struggle to synchronise visual cues with auditory to support NWR leading to minimal change in production accuracy in comparison to audio only presentation. In contrast, typically developing controls would be better able to integrate predominantly auditory information such as voicing and manner with visual cues including articulator placement (Summerfield, 1992), allowing more accurate repetition. The primary hypothesis therefore contained two parts as summarised below:

Hypothesis 1

a) Children with LLE would show lower repetition scores than the NLE group for both Audio (A) and Audiovisual (AV) presentation of non-words,
b) The LLE children would respond less to visual cues than the NLE group during the repetition of non-words as demonstrated by a lower AV over A benefit.

As NWR proficiency is considered to be dependent on multiple factors it was proposed that, in children with LLE, weak audiovisual synchrony would interact with difficulties in other key areas believed to influence phonological processing, namely memory load as measured by syllable length and articulation demand in relation to non-word phoneme acquisition age. This was investigated using the following supplementary hypotheses:

Hypothesis 2:

a) The LLE group would show lower NWR scores than the NLE group for both A and AV presentation as syllable length increased,
b) AV over A benefit would be smaller for the LLE group than the NLE group at each syllable length.

Hypothesis 3:

a) The LLE group would show lower NWR scores than the NLE group for both A and AV presentation as phonemes acquisition age increased
b) AV over A benefit would be smaller for the LLE group than the NLE group at each phoneme acquisition level.

Method

All procedures for the study gained ethical approval from the Language and Communication Science Proportionate Review Committee at City University London.

Participants

Children were selected from a larger NWR investigation carried out at the Science Museum in London across 30 days and in affiliation with City University London. The research was publicised on the Museum website, as well as through leaflets and information signs at the Museum front entrance and in the area where testing took place.

Children were recruited if they or their parent/guardian expressed an interest in participating. The parent/guardian was asked to complete a demographic questionnaire (see Appendix A) and children were considered for the LLE group if:

- first word emerged after 12 months;
- English was a native language, this could be with mono- or multi-lingual presentation
- no language related difficulties were identified[1]

Whilst previous studies have tended to classify LLE based on a delay in combining words as identified retrospectively by parents (cf. Rice, Taylor, & Zubrick, 2008; Reilly et al., 2009), this study focused on first word emergence instead. This allowed patterns of delay to be determined from an earlier stage therefore guiding clinical practice more effectively. In addition, a strong association has previously been found between parental report of delay in first word use and reported delay in use of word-combinations (Preston et al., 2010).

A matched-pair control group was formed using children from the same larger study whose first words emerged at around 12 months or younger, i.e. normal language emergence (NLE). The final sample consisted of 25 children in the LLE group and 25 in the NLE group all aged between 4 years 0 months and 6 years 11 months (see Table 1).

The lower age limit was to differentiate between transient and persistent language impairment as children with difficulties resolving before school-age are generally considered late-talkers whereas issues beyond the preschool period are considered more indicative of a chronic language impairment (Paul and Roth, 2011). The upper boundary was selected given that studies have indicated weaker correspondence between NWR and linguistic skills in children over 6 (Gathercole, 2006; Chiat & Roy, 2013), furthermore, the cut-off of 83 months

[1] Although children were not excluded on the basis of sensory, cognitive or physical difficulties, the final sample did not contain any participants with overt needs in these areas.

allowed normality of distribution of NWR scores to be retained enabling more robust statistical analysis.

Twenty children were categorised by their parent's/guardian's as having language which emerged around 18 months, the remaining five children were identified as talking from approximately 24 months. Independent sample t-tests showed that there was not a significant difference in either audio ($t(23) = 1.27$, $p = .22$) or audiovisual scores ($t(23) = .47$, $p = .64$) between the two groups, the decision was therefore made to made to cluster all the late-talking children into one group as a larger sample size would increase the strength of statistical findings. Appendix B details the complete selection process.

It was possible to form fully matched pairs for gender and family history. All but one pair were matched on presence of a language in addition to English. It was more difficult to find exact matches in terms of age however an independent samples *t*-test confirmed that the final two groups did not differ significantly on this variable ($t(48) = .11$, $p = .91$). Appendix C contains full demographic details for each matched pair.

Table 1

Summary of Participant Characteristics

Demographic information	LLE n = 25	NLE n = 25	Total N = 50
Male			
n (%)	7 (28%)	7 (28%)	14 (28%)
Female			
n (%)	18 (72%)	18 (72%)	36 (72%)
Family history			
n (%)	3 (12%)	3 (12%)	6 (12%)
Monolingual			
n (%)	17 (68%)	18 (72%)	35 (70%)
Multilingual			
n (%)	8 (32%)	7 (28%)	15 (30%)
Age in months			
Mean (SD)	68.36 (10.80)	68.68 (9.70)	68.52 (10.16)

Note. LLE = Late Language Emergence; NLE = Normal Language Emergence.

Stimuli

Forty-eight pseudo words were created to form the stimulus set. Syllable number ranged from 2-5 with 12 non-words at each length. Within each length the non-words formed three equal-sized age of acquisition (AoA) categories: early (sounds typically acquired around 1-3 years), middle (3-6 ½ years) and late (5-7 ½ years) (Shriberg, 1993). Half the non-words in each syllable length had standard English stress pattern and half had atypical stress. The final stimulus set contained an even distribution of syllable length, sound acquisition level and stress pattern.

To ensure consistency of presentation, the stimuli were pre-recorded by a native English speaker trained in phonetics. In order to appeal to the paediatric population, the repetition task revolved around the theme of learning an ‘alien language’ with the speaker using face makeup and a wig to represent an ‘alien’ (see Figure 1). To minimise distraction only the speaker’s face was visible on the recording which was digitised at a rate of 25 frames per second, audio was captured simultaneously using an external microphone.

For each non-word, an audio only and audiovisual version was created providing 96 stimuli items in total. This was used to form two non-word data sets; the first set labelled ‘Alien 1’ consisted of half the non-words within each syllable length, AoA and stress pattern being audio only and the remainder audiovisual, the second set ‘Alien 2’ reversed the pattern so that the audio non-words were now audiovisual and vice-versa. Data collectors were asked to alternate between the two data sets for each subsequent participant and this ensured that data was obtained for all 48 non-words in both audio and audiovisual formats. The control group only included children who were introduced to the same data set as the LLE group.

The non-words were presented in three blocks: the first containing 2-3 syllable non-words, the second 3-4 and the final 4-5. The blocks were always presented in the same order moving from 2 to 5 syllables, however within each block stimuli was presented in a pseudo-random order in terms of presentation mode (audio/audio-visual), syllable length, age of acquisition (early, middle, late) and stress pattern (usual/unusual). Appendix E provides a detailed breakdown of the stimuli set and order of presentation.

		
Child hears and sees a recording of a non-word spoken by an ‘alien’ character	Child’s repetition attempt entered by the data-collector (syllables correct) with the audio recorded for follow-up analysis	Child receives response-contingent feedback onscreen and via the data-collector

Fig 1. Demonstration of an audiovisual non-word repetition trial

Procedure

The testing environment was a semi-partitioned section of a large open-access exhibition space within the Science Museum. Participants were presented with the pre-recorded non-word data sets from a Dell Latitude laptop with a 15 x 13 cm screen. For consistency each participant was seated facing the laptop from a distance of approximately 45 cm. Twenty four of the non-words in each set involved audio only presentation (hearing a recording of the speaker whilst the screen stayed blank) and the remaining 24 audiovisual (both hearing and seeing the speaker).

The noise level for the testing environment was calculated to be approximately equal to the signal at 65dB, however audio stimuli were presented through over-ear headphones which served to minimise the influence of external noise by increasing the signal to noise ratio to above zero. All participant responses were recorded using a small desk-top microphone facing towards them and positioned directly in front of the laptop.

Data collectors used headphone splitters with an earpiece in one ear to allow both the non-word stimuli and participant's repetition attempts to be heard. Following each repetition attempt, the data collector entered the number of syllables produced correctly using the laptop keyboard. This triggered positive feedback contingent to the participant's response to appear on the screen which was further reinforced by the data collector, following this, the next non-word stimuli was presented (see Figure 1). A break was introduced after each block, the length of which was based on participant need as judged subjectively by the data collector but did not exceed 1 minute durations. On average the entire task lasted around 10 minutes.

During the study the stimuli was introduced to the child whilst the parent/guardian completed the demographic questionnaire, as a result the data collector was blind to participant details during the testing process. Participant data was excluded if attempts were not made for each stimulus or if insufficient demographic information was provided.

Analysis

Scoring. Each NWR attempt was scored according to the number of syllables produced correctly. A syllable was considered correct if all of the consonants within the syllable were articulated accurately and in the right position. Omissions and substitutions were considered incorrect however dialectal differences including commonly accepted allophonic variations as well as mild vowel and stress distortions were allowed if manifested consistently by the participant.

NWR attempts were initially scored manually into the laptop by the data collector immediately following participant production. All the data was then re-checked by the same data collector at a later date in a quiet non-distracting environment using the audio recordings; 10% of the data was further checked by a fellow data collector and adequate inter-rater reliability confirmed using Cohen's Kappa test (Cronbach's $\alpha = 0.94$).

Statistical tests. All statistics involved parametric tests that were carried out using IBM SPSS© Statistics 21.00, with significance set at $p < .05$. Normality and sphericity of distributions were either confirmed or relevant statistical adjustments made as detailed in the results section. Mixed-model analysis of variance (ANOVA) tests were used to identify differences in mean scores for between- and within-group variables. Post hoc tests, consisting of Bonferroni corrected multiple comparison *t*-tests and planned paired-sample *t*-tests were used to identify the source of significant differences if their presence was indicated during the data analysis. Effect sizes were reported where relevant using partial-eta squared (η_p^2) for ANOVA and Cohen's *d* for *t*-tests as defined below:

- η_p^2**:** < .01 = negligible effect; .01 = small; .06 = medium; > .14 = large (Dattalo, 2008)
- *d:* .2 = small effect; .5 = medium, .8 = large (Cohen, 1988). As SPPS does not calculate Cohen's *d* an online tool was used: http://www.uccs.edu/lbecker/index.html

Results

For each participant, the number of syllables repeated correctly was calculated as a percentage of the total number of syllables presented. A 100% score indicated that no errors were made in the repetition of any syllable within all of the non-words presented, whereas a score of 50% indicated that half the syllables in all of the non-words presented were repeated inaccurately. Scores were then summed for each group to allow mean and standard deviation values to be calculated (LLE: M = 76.33, SD = 11.76; NLE: M = 81.12, SD = 12.90).

Normality of distribution was confirmed for both the LLE (z^{skew} = 1.10; $z^{kurtosis}$ = 1.00) and the NLE group (z^{skew} = 1.90; $z^{kurtosis}$ = .23) thus allowing parametric tests to be run on the data. Where sphericity could not be assumed Greenhouse-Geisser adjustments were used.

Hypothesis 1

a) The LLE group would show lower repetition scores than the NLE group for both Audio (A) and Audiovisual (AV) presentation of non-words,

b) The LLE group would show a smaller AV over A benefit than the NLE group.

A “2 (Group: LLE, NLE) x 2 (Presentation: A, AV) mixed ANOVA” was used. Table 2 provides a summary of the descriptive statistics. A main effect was found for Presentation ($F(1, 48) = 7.45$, $p = .01$, $\eta_p^2 = .13$, large effect size). There was not a significant effect of Group ($p = .18$) nor a Group x Presentation interaction ($p = .10$), however visual representation of the data appeared to indicate a greater disparity between A and AV results for the LLE than the NLE group (see Fig 2). The decision was therefore made to split the data by Group and run paired *t*-tests using A and AV scores. Findings confirmed that although intergroup differences were not sufficient to drive an interaction effect, AV scores were significantly higher than A scores for the LLE group (M diff = 4.10, SD = 6.17; $t(24) = 3.32$, $p = .003$, $d = .34$, small to medium effect size) but not the NLE group ($p = .50$).

Table 2

Scores for Audio only and Audiovisual Presentation of Non-Word Stimuli

Presentation	Group	M	SD	N
A	LLE	74.38	13.52	25
AV	LLE	78.48	10.65	25
A	NLE	80.67	12.66	25
AV	NLE	81.62	14.01	25

Note. A = Audio; AV = Audiovisual. LLE = Late Language Emergence; NLE = Normal Language Emergence.

The hypothesis of LLE children having lower NWR scores than the NLE group was not supported through statistical testing, as although the LLE group did achieve lower scores than the NLE group with both A and AV presentation differences did not reach significance.

The hypothesis of LLE children showing lower AV over A benefit than the NLE group was also not supported, in fact the LLE children appeared to benefit to a greater extent.

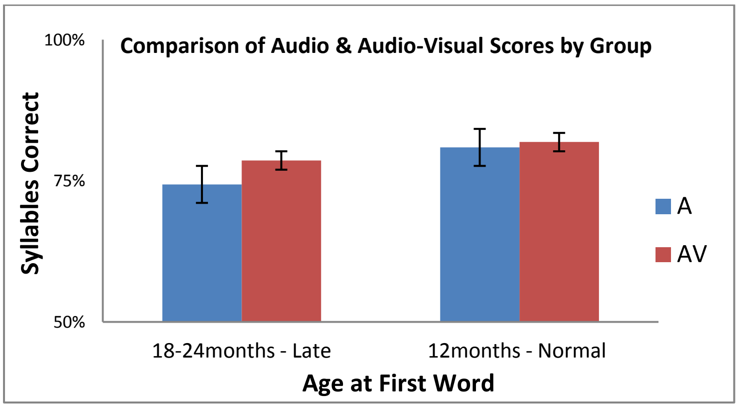


Fig 2. Mean percentage of syllables correct and standard errors for Late and Normal Language Emergence Groups. A = Audio only presentation; AV = Audiovisual presentation

Hypothesis 2:

a) The LLE group would show lower NWR scores than the NLE group for both A and AV presentation as syllable length increased,

b) The LLE group would show a smaller AV over A benefit than the NLE group at each syllable length.

A “2 (Group: LLE, NLE) x 2 (Presentation: A, AV) x 4 (Syllable Length: 2Syll, 3Syll, 4Syll, 5Syll) mixed ANOVA” was used. Table 3 summarises the descriptive statistics.

A main effect was found for Syllable Length ($F(2.32, 111.45) = 38.17$, $p < .001$, ηp2 = .44, large effect size) and an interaction effect for Group x Syllable Length ($F(2.32, 111.45) = 3.80$, $p = .02$, $\eta_p^2 = .07$, medium effect size). There was not a Presentation x Syllable Length interaction ($p = .14$) nor a Group x Presentation x Syllable Length three way interaction ($p = .58$).

Table 3

Scores for Audio only and Audiovisual Presentation of Non-Word Stimuli of Increasing Syllable Length

Condition	Group	M	SD	N
A2Syll	LLE	76.67	17.35	25
	NLE	81.00	17.10	25
AV2Syll	LLE	85.67	10.35	25
	NLE	84.33	13.46	25
A3Syll	LLE	84.67	11.82	25
	NLE	86.44	10.77	25
AV3Syll	LLE	88.44	9.20	25
	NLE	87.11	12.50	25
A4Syll	LLE	79.33	16.20	25
	NLE	82.17	13.20	25
AV4Syll	LLE	82.50	13.45	25
	NLE	85.33	16.58	25
A5Syll	LLE	62.80	18.27	25
	NLE	75.73	18.22	25
AV5Syll	LLE	66.40	16.50	25
	NLE	74.27	17.41	25

Note. A = Audio; AV = Audiovisual. Syll = Syllable Number. LLE = Late Language Emergence; NLE = Normal Language Emergence.

To further investigate the Group x Syllable Length interaction, comparison were made between the two groups at each syllable length using individual *t*-tests. The difference in 5 syllable scores between the two groups reached a *p*-value of 0.03 (NLE: M = 75.00, SD = 17.18; LLE: M = 64.60: SD = 16.23; $t(48)$ = 2.20, d = 0.62 - medium effect). This value lies within the standard $p < .05$ level of significance however falls short of the more stringent Bonferroni corrected multiple *t*-tests value of $p < 0.01$. No other between-group differences achieved a *p*-value below 0.05 indicating lack of significance.

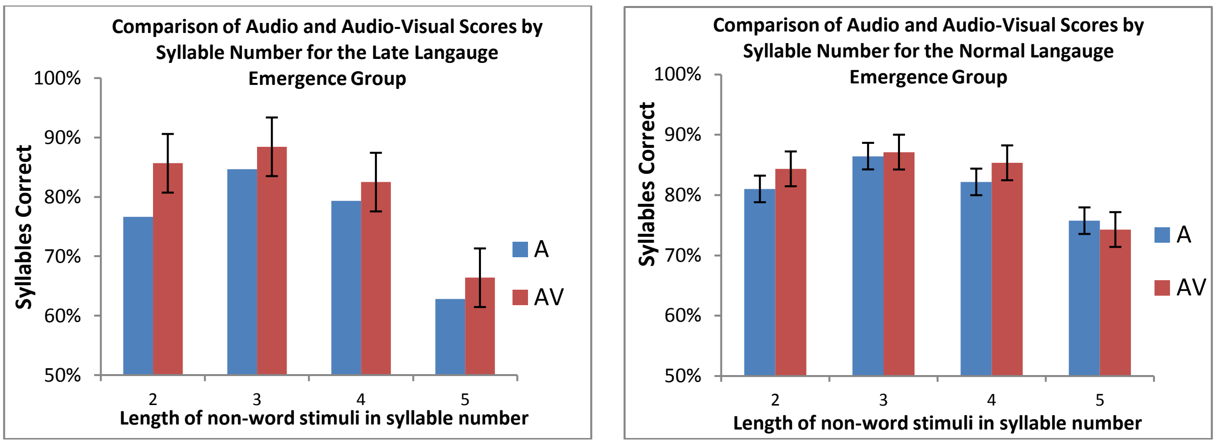


Fig 3. Mean percentage syllables correct and standard errors at each Syllable Length. Separate charts for Late (18-24months) & Normal (12 month) Language Emergence Groups. A = Audio; AV = Audiovisual presentation.

The hypothesis that the LLE group would show lower scores with both A and AV presentation than the NLE group as non-word syllable length increased was supported. Both groups showed a reduction in performance as syllable length increased regardless of presentation except with 2 syllable non-words. The gap between LLE and NLE widened as syllable length increased with the difference reaching significance at the 5 syllable length.

The hypothesis of LLE children showing a smaller difference in AV than A scores at each syllable length in comparison to the NLE group was not supported. The pattern of LLE children demonstrating greater benefit than the NLE group from visual cuing did not alter as syllable number increased. Interestingly the LLE children appeared to benefit most from the introduction of visual cueing with 2 syllable length non-words as can be seen in Figure 3.

Hypothesis 3:

a) The LLE group would show lower NWR scores than the NLE group for both A and AV presentation as phonemes acquisition age increased
b) The LLE group would show a smaller AV over A benefit than the NLE group at each acquisition level.

A "2 (Group: LLE, NLE) x 2 (Presentation: A, AV) x 3 (AoA: Early, Mid, Late) mixed ANOVA" was used. Table 5 provides a summary the descriptive statistics. There was a main effect of AoA ($F(1.77, 85.17) = 19.61$, $p <.001$, $\eta_p^2 = .29$ - large effect size) but no interaction effect of either Group x AoA ($p = .61$) or Presentation x AOA ($p = .19$); a Group x Presentation x AoA three way interaction was also not found ($p = .10$).

Post-hoc Bonferroni corrected *t*-tests showed that Early AoA scores were significantly higher than Mid (*M* diff = 8.3, *SE* = 1.2, $p < .001$, $d = 0.67$, medium-large effect size) and Late AoA responses (*M* diff = 7.5, *SE* = 1.6, $p < .001$, $d = 0.53$, medium effect size); the difference between Mid and Late responses was not significant ($p = 1.0$).

Table 5

Scores for Audio and Audiovisual Presentation of Non-Word Stimuli of Increasing Phoneme Acquisition Level

Condition	Group	M	SD	N
A.Early	LLE	80.43	12.55	25
	NLE	85.86	10.07	25
AV.Early	LLE	84.29	10.46	25
	NLE	85.29	14.00	25
A.Mid	LLE	69.14	14.39	25
	NLE	77.71	15.75	25
AV.Mid	LLE	77.00	12.07	25
	NLE	79.00	13.84	25
A.Late	LLE	73.00	18.47	25
	NLE	78.28	18.12	25
AV.Late	LLE	74.14	15.27	25
	NLE	80.57	17.01	25

Note. A = Audio; AV = Audiovisual. Early, Mid, Late = Age of Acquisition. LLE = Late Language Emergence; NLE = Normal Language Emergence.

The hypothesis that LLE children would show lower scores than the NLE group as AoA increased was not supported. Both groups showed significantly lower scores for Mid and Late acquired sounds in comparison to Early sounds regardless of presentation and whilst there was a trend for LLE scores to be lower than NLE differences did not reach significance.

The hypothesis of LLE children showing a smaller improvement in AV than A scores was also not supported. As shown in Figure 4, the LLE group appeared to benefit more from visual cuing than the NLE children with sounds that are acquired in early- or mid- childhood, however between-group differences were not sufficient to drive an interaction effect.

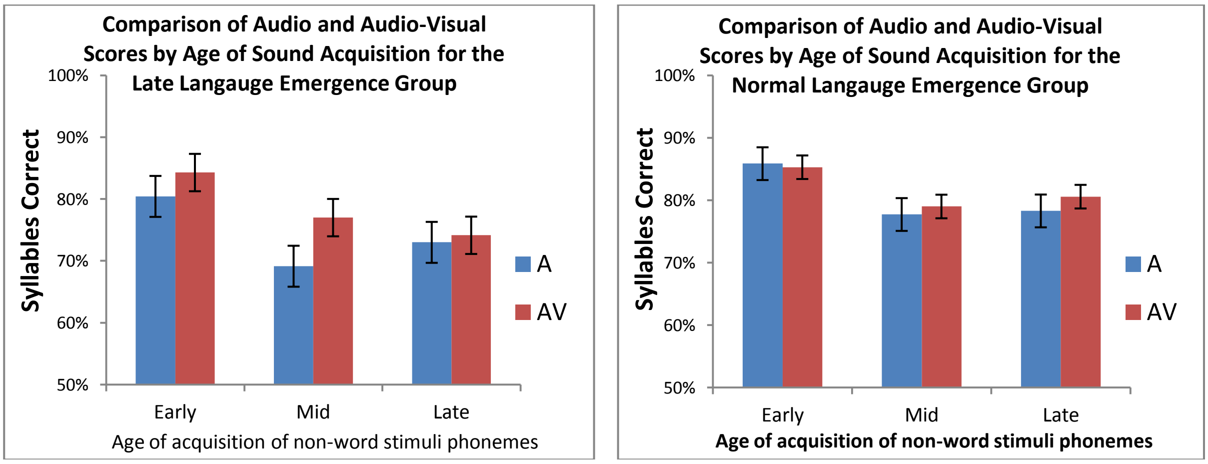


Fig 4. Mean percentage of syllables correct and standard errors at each Age of Acquisition Level. Separate charts for Late (18-24months) and Normal (12months) Language Emergence Groups. A = Audio; AV = Audiovisual.

Discussion

The primary aim of this study was to investigate the impact of supplementing auditory verbal information with visual speech cues on NWR in children with delayed language onset.

Several studies have suggested that difficulties with audiovisual speech perception in children with a persistent language impairment may be restricting the extent to which visual cues support auditory signals during verbal processing (Meronen et al., 2013; Norrix et al., 2007; Pons et al., 2012). Given that profiles for persistent and resolved language impairment may represent two ends of a spectrum of severity (Rescorla, 2002), it was hypothesised that children who began talking late but appeared to have resolved their language needs would also gain less benefit from visual reinforcement of auditory information. Comparison of audio and audiovisual performance on NWR tasks was used to investigate the hypothesis. This was because manipulation of non-word speech has been shown to be a particular area of difficulty for children with early language learning difficulties regardless of subsequent progress made with language skills (Bishop, et al., 1996; Conti-Ramsden et al., 2001).

Previous studies have proposed that NWR difficulties represent deficits in a range of underlying speech processing skills in particular memory load and articulation competency (Estes, 2007). The relationship between these two areas and response to audiovisual cuing was therefore also considered in order to learn more about the nature of factors influencing language learning. Findings from each area studied are considered below.

Audio vs. audiovisual performance on NWR tasks

The number of non-word syllables repeated correctly when presented as audio-only or audiovisual was used to determine if there was a difference in performance between children with late and normal language emergence. As expected, the LLE children scored lower on the NWR tasks than the NLE group, however the difference failed to reach significance.

In addition, contrary to the hypothesis of late-talkers showing less sensitivity to visual cuing, they actually showed a greater improvement in scores than the control group, with the difference between audio and audiovisual performance reaching significance.

Relationship between syllable length and audiovisual cuing on NWR tasks

Both the typical and delayed language onset groups showed a general decline in audio and audiovisual scores as syllable length increased. The LLE children, however, appeared to underperform to a greater extent than the NLE group as non-word length grew, with differences reaching significance at the 5 syllable level.

The pattern of 2 syllable NWR scores is worth noting for both groups as performance was weaker than for longer syllable strings despite the lower demand on memory load. This could have been due to the repetition tasks starting at this syllable length for all children, therefore focus may have been on task familiarisation rather than production accuracy.

The markedly greater reduction in performance of the LLE group as syllable length increased indicates a more restricted recall capacity in the children with late language acquisition. This reflects findings from Bishop et al. (1996) who also found greater disassociation in NWR performance between late-talking children and typical controls as syllable length increased.

Once again the LLE children showed greater benefit from the introduction of visual cues than the NLE group, this was particularly evident at the 2 syllable level where the difference between audio and audiovisual scores was the most prominent.

Relationship between phonemic demand and audio-visual cuing on NWR tasks

Both groups showed a significant reduction in audio and audiovisual performance when NWR scores for early acquired sounds was compared with scores for sounds mastered in mid or late childhood. This is unsurprising as the earlier emerging sounds are considered less challenging in terms of perception and production. Whilst the NLE group consistently outperformed the LLE children, differences failed to reach significance.

The LLE group appeared to have higher audiovisual than audio-only scores with early- and mid- acquired sounds, whereas the NLE children showed minimal benefit from visual cues.

Conclusion

The findings from this study, with regards to the influence of visual cues on speech perception in children with apparently resolved language difficulties, failed to replicate that of studies involving children with persistent language issues. Whereas children with ongoing language impairments fail to derive benefit from the introduction of visual cues during verbal processing (Norrix et al., 2007; Pons et al., 2012), the children in this study demonstrated consistently improved performance as measured using NWR tasks.

These outcomes can be considered within the context of current theories on audiovisual speech perception. It is believed that in early infancy, verbal processing involves watching a speaker’s mouth and relating this to what is heard in order to store word forms as part of a phonological-based learning mechanism. Once audiovisual synchrony is achieved, visual priority shifts to semantic and social rather than phonological cues in order to map form to meaning. As a result focus shifts from the speaker’s mouth to other visual sources such as eye gaze whilst phonological information is derived primarily through the auditory channel (Lewkowicz & Hansen-Tift, 2012). Eventually as phonological forms become embedded, visual speech cues become redundant and are only called upon when the auditory signal is degraded, so an adult is typically only likely to watch a speaker’s mouth if they are unable to hear them well (Campbell, 2006). As minimal benefit from augmenting auditory speech cues with visual during verbal perception tasks appears characteristic of persistent language impairments, it has been argued that the verbal deficits may be linked to underlying audiovisual processing issues which are present from infancy (Pons et al., 2012). In comparison, given that the LLE group in this study did benefit from visual cues during speech processing, it could be argued that children with non-persistent language difficulties do reach the early audiovisual integration stage of speech processing.

However, whilst the profile of children with resolved language difficulties did not mirror that of children with persistent needs, it also did not correspond to children with typical language emergence. In the majority of cases, the late-talking children in this study demonstrated greater benefit from visual information, though not greater repetition accuracy, than normal language development controls. It could be argued that the children with typical language onset obtained sufficient information to process the non-words through audio means alone making the visual cues redundant. The fact that over half the control group presented with average NWR scores of 85% correct or greater with audio only presentation reinforces this assumption. An alternative explanation is that the typically developing group had progressed from an audiovisual to an audio dominant speech processing system but were still developing their ability to augment auditory information using visual cues when verbal processing demands increased, a skill which does not fully develop until adulthood (Lewkowicz & Flom, 2013).

The fact that the increased benefit of visual cues for the late-talking children in comparison to the children with normal language onset did not translate into comparatively higher NWR scores is worth further consideration. Although findings rarely reached significance, the LLE children consistently underperformed during NWR tasks in comparison to the LLE group. This adds to the body of studies highlighting NWR skills as a marker of language learning competency even when there is no longer an identified communication impairment (Bishop, et al., 1996; Conti-Ramsden et al., 2001). It is also in line with numerous studies which have found that whilst late-talking children fall within the average range on verbal assessments which have been standardised on populations with typical language development, there is a tendency for performance to cluster at the low average point (Rescorla, 2011).

Furthermore, in keeping with findings by Bishop et al. (1996), this study found that the children with delayed language emergence struggled considerably more than typically developing peers with speech manipulation as memory load, in the form of syllable length, increased. Also, as previously argued by Estes et al., (2007), the articulatory complexity of the sound stimuli was shown in this study to have a more detrimental effect on accuracy of production in late-talkers than children with typical language emergence despite neither group having overt articulation difficulties. This indicates that whilst visual speech processing skills may be intact for children with late language onset they may still struggle with other aspects of verbal processing including memory load and phonemic competency thereby leading to an immature language processing profile which is overly dependent on visual verbal cues.

A useful insight into the influence of visual cues on memory load and phonemic demand can be gained from this study. Visual cues appeared to be most beneficial to late-talking children when memory and articulation demands were relatively low, i.e. with non-words consisting of 2 syllables and sounds acquired in early and middle childhood. It could be that with non-words of longer length and greater phonemic complexity the processing and production demands superseded the benefits gained from the visual cues leading to a smaller difference in performance compared to audio only presentation.

In terms of follow-up, useful insight could be gained from directly comparing audio and audiovisual manipulation of non-words in children with typical, resolved and persistent language difficulties. Differences in speech processing styles may provide an indication of underlying deficit as well as informing needs-specific management decisions.

Clinical Implications

This study has demonstrated that a delay in language emergence of just 6-12 months may not only reduce the efficiency of verbal processing but could also be associated with an over-dependency on visual speech cues in school-aged children. The functional implications of late-talking are well documented and include low-average language skills as well as lower academic attainment, wellbeing and independence than in peers with typical language emergence (Shevell, 2005; Snowling, et al., 2001).

As there is currently no accurate way of predicting exactly which children will have late language onset or who from this group will have persistent difficulties (Reilly, et al., 2010), it is imperative that there is salient and coordinated exposure to both audio and visual components of language from infancy onwards. For school-aged children this audiovisual access will be of particular importance in educational settings given the implicit role that language processing plays in many aspects of learning within the classroom environment (Gathercole, Durling, Evans, Jeffcock, & Stone, 2008).

Limitations

A particular limitation of this study was the above expected performance of not only the typically developing but also the late-talking group where average syllable correct scores did not fall below 50%. The likely source of this was the recruitment method used; the investigation was based at the Science Museum in London and families volunteered to participate whilst onsite. Annual UK visitor information from the Science Museum has indicated low representation from what is considered “lower socioeconomic backgrounds” (Science Museum Group, 2013). Given that socioeconomic status is thought to influence language onset (Roy & Chiat, 2012), this under-representation may have led to higher NWR performance results. Repeating the study with a more socially diverse population as well as utilising more demanding NWR tasks may provide greater insight in the future.

A further limitation was that the proportion of females expressing an interest in participation and therefore representing the final LLE group was far higher than males (72:38). This distribution did not reflect the late-talking population where numbers have been found to be far greater amongst boys than girls (cf. McLeod & Harrison, 2009) and may have skewed the results.

Another area to consider is the stimuli set used in the study. The non-words were taken from a larger investigation of audiovisual speech perception where familiarity of sound sequences (phonotactic probability) and presence of consonant clusters were identified retrospectively. This resulted in an uneven distribution across the stimuli set, preventing reliable analysis of impact on performance despite the possibility of implications on NWR performance in late-talking children (Bishop et al., 1996; MacRoy-Higgins, 2013). Previous studies have shown strong parallels between the influence of consonant clusters and later acquired phonemes (Estes, 2007) and the latter was considered in this study, however there was not a similar overlap for phonotactic probability within this study making it an important area for consideration in subsequent research projects.

References

Bishop, D. V. (2006). What causes specific language impairment in children? *Current directions in psychological science*, *15*(5), 217-221.

Bishop, D. V., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry*, *37*(4), 391-403.

Boelhouwer, W. (2011). *Phonotactic Probability and Neighborhood Density in Nonword Learning in Typical Language Development and Specific Language Impairment.* [Doctoral dissertation, University of Amsterdam] Retrieved at: http://dare.uva.nl/document/343192 [Accessed 26 May. 2014]

Botting, N., & Conti-Ramsden, G. (2000). Social and behavioural difficulties in children with language impairment. *Child Language Teaching and Therapy*, *16*(2), 105-120.

Broomfield, J., & Dodd, B. (2004). Children with speech and language disability: Caseload characteristics. *International Journal of Language & Communication Disorders, 39*(3), 303-324.

Campbell, R. (2008). The processing of audio-visual speech: Empirical and neural bases. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 363*(1493), 1001-1010. doi:723644J2W3155U20 [pii]

Chiat, S. (2001). Mapping theories of developmental language impairment: Premises, predictions and evidence. *Language and Cognitive Processes*, *16*(2-3), 113-142.

Chiat, S., & Roy, P. (2013). Early predictors of language and social communication impairments at ages 9–11 years: A follow-up study of early-referred children. *Journal of Speech, Language, and Hearing Research, 56*(6), 1824-1836.

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders, 43*(1), 1-40.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, 42*(6), 741-748.

Darley, F. L., & Winitz, H. (1961). Age of first word: Review of research. *Journal of Speech and Hearing Disorders, 26*(3), 272.

Darrah, J., Hodge, M., Magill-Evans, J., & Kembhavi, G. (2003). Stability of serial assessments of motor and communication abilities in typically developing infants—implications for screening. *Early Human Development, 72*(2), 97-110.

Dattalo, P. (2008). *Determining sample size: Balancing power, precision, and practicality* Oxford University Press, USA.

Davies, R., Kidd, E., & Lander, K. (2009). Investigating the psycholinguistic correlates of speechreading in preschool age children. *International Journal of Language & Communication Disorders, 44*(2), 164-174.

Dodd, B. (1979). Lip reading in infants: Attention to speech presented in-and out-of-synchrony. *Cognitive Psychology, 11*(4), 478-484.

Domsch, C., Richels, C., Saldana, M., Coleman, C., Wimberly, C., & Maxwell, L. (2012). Narrative skill and syntactic complexity in school-age children with and without late language emergence. *International Journal of Language & Communication Disorders, 47*(2), 197-207.

Estes, K. G., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 50*(1), 177-195.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27*(04), 513-543.

Gathercole, S. E., Durling, E., Evans, M., Jeffcock, S., & Stone, S. (2008). Working memory abilities and children's performance in laboratory analogues of classroom activities. *Applied Cognitive Psychology, 22*(8), 1019-1037.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29*(3), 336-360.

Henrichs, J., Rescorla, L., Schenk, J. J., Schmidt, H. G., Jaddoe, V. W., Hofman, A., . . . Tiemeier, H. (2011). Examining continuity of early expressive vocabulary development: The generation R study. *Journal of Speech, Language, and Hearing Research, 54*(3), 854-869.

Hosozawa, M., Tanaka, K., Shimizu, T., Nakano, T., & Kitazawa, S. (2012). How children with specific language impairment view social situations: An eye tracking study. *Pediatrics, 129*(6), e1453-60. doi:10.1542/peds.2011-2278 [doi]

Jusczyk, P. W., & Luce, P. A. (1994). Infants′ sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, *33*(5), 630-645.

Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huotilainen, M., & Näätänen, R. (2002). Maturation of the auditory event-related potentials during the first year of life. *Neuroreport*, *13*(1), 47-51.

Lewkowicz, D. J., & Flom, R. (2013). The audiovisual temporal binding window narrows in early childhood. *Child Development,*

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(5), 1431-1436. doi:10.1073/pnas.1114783109 [doi]

MacRoy-Higgins, M., Schwartz, R. G., Shafer, V. L., & Marton, K. (2013). Influence of phonotactic probability/neighbourhood density on lexical learning in late talkers. *International Journal of Language & Communication Disorders, 48*(2), 188-199.

Marinis, T. (2011). On the nature and cause of specific language impairment: A view from sentence processing and infant research. *Lingua, 121*(3), 463-475.

McLeod, S., & Harrison, L. J. (2009). Epidemiology of speech and language impairment in a nationally representative sample of 4-to 5-year-old children. *Journal of Speech, Language, and Hearing Research*, *52*(5), 1213-1229.

Meronen, A., Tiippana, K., Westerholm, J., & Ahonen, T. (2013). Audiovisual speech perception in children with developmental language disorder in degraded listening conditions. *Journal of Speech, Language, and Hearing Research*, *56*(1), 211-221.

Moore, M. W., Tompkins, C. A., & Dollaghan, C. A. (2010). Manipulating articulatory demands in non-word repetition: A'late-8'non-word repetition task. *Clinical Linguistics & Phonetics, 24*(12), 997-1008.

Norrix, L. W., Plante, E., Vance, R., & Boliek, C. A. (2007). Auditory-visual integration for speech by children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 50*(6), 1639-1651.

Paul, D., & Roth, F. P. (2011). Guiding principles and clinical applications for speech-language pathology practice in early intervention. *Language, Speech, and Hearing Services in Schools, 42*(3), 320-330.

Pons, F., Andreu, L., Sanz-Torrent, M., Buil-Legaz, L., & Lewkowicz, D. J. (2013). Perception of audio-visual speech synchrony in Spanish-speaking children with and without specific language impairment. *J.Child Lang, 40*, 687-700.

Preston, J. L., Frost, S. J., Mencl, W. E., Fulbright, R. K., Landi, N., Grigorenko, E., . . . Pugh, K. R. (2010). Early and late talkers: School-age language, literacy and neurolinguistic differences. *Brain : A Journal of Neurology, 133*(Pt 8), 2185-2195. doi:10.1093/brain/awq163 [doi]

Reilly, S., Bavin, E. L., Bretherton, L., Conway, L., Eadie, P., Cini, E., . . . Wake, M. (2009). The early language in Victoria study (ELVS): A prospective, longitudinal study of communication skills and expressive vocabulary development at 8, 12 and 24 months. *International Journal of Speech and Language Pathology, 11*(5), 344-357.

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., . . . Bretherton, L. (2010). Predicting language outcomes at 4 years of age: Findings from early language in Victoria study. *Pediatrics, 126*(6), e1530-7. doi:10.1542/peds.2010-0254 [doi]

Rescorla, L. (2002). Language and reading outcomes to age 9 in late-talking toddlers. *Journal of Speech, Language & Hearing Research, 45*(2)

Rescorla, L. (2011). Late talkers: Do good predictors of outcome exist? *Developmental Disabilities Research Reviews, 17*(2), 141-150.

Rice, M. L., Taylor, C. L., & Zubrick, S. R. (2008). Language outcomes of 7-year-old children with or without a history of late language emergence at 24 months. *Journal of Speech, Language, and Hearing Research, 51*(2), 394-407.

Roy, P., & Chiat, S. (2012). Teasing apart disadvantage from disorder: The case of poor language. In C. Marshall (Ed) *Current Issues in Developmental Psychology*, (pp. 125-150)

Sahlen, B., Reuterskiold-Wagner, C., Nettelbladt, U., & Radeborg, K. (1999). Non-word repetition in children with language impairment-pitfalls and possibilities. *International Journal of Language & Communication Disorders, 34*(3), 337-352.

Science Museum Group, (2013). *Science museum group written evidence to the culture media and sport select committee inquiry into the future of the science museum group.* [PDF] Retrieved from: http://www.sciencemuseum.org.uk/documents/Select-committee-Written-evidence-27-June-2013.pdf [Accessed 26 May. 2014]

Shevell, M. I., Majnemer, A., Webster, R. I., Platt, R. W., & Birnbaum, R. (2005). Outcomes at school age of preschool children with developmental language impairment. *Pediatric Neurology, 32*(4), 264-269.

Shriberg, L. D. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech, Language, and Hearing Research, 36*(1), 105-140.

Snowling, M. J., Adams, J. W., Bishop, D. & Stothard, S. E. (2001). Educational attainments of school leavers with a preschool history of speech-language impairments. *International Journal of Language & Communication Disorders, 36*(2), 173-183.

Snowling, M. J., Bishop, D., Stothard, S. E., Chipchase, B., & Kaplan, C. (2006). Psychosocial outcomes at 15 years of children with a preschool history of speech-language impairment. *Journal of Child Psychology and Psychiatry, 47*(8), 759-765.

Summerfield, Q. (1992). Lipreading and audio-visual speech perception. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 335*(1273), 71-78.

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 39*(6), 1284-1294.

Young, G. S., Merin, N., Rogers, S. J., & Ozonoff, S. (2009). Gaze behavior and affect at 6 months: Predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. *Developmental Science, 12*(5), 798-814.

Appendices

Appendix A

Demographic questionnaire – completed by parent/guardian of child participating in the study

CITY UNIVERSITY LONDON

EST 1894

Pt.........

Language Questionnaire

This questionnaire asks about your child's language. When doing any scientific research it is important to know as much about volunteers as possible. Here it is particularly valuable to find out about children's language background. *Please provide the following information about your child:*

1. Date of birth ☐

2. Gender Male ☐ Female ☐

3. Is English your child's native language? Yes ☐ No ☐

If English is not their native language:

3.a How long have they been learning English for? ☐ Years

3.b What is(are) their native language(s)? ....

3.c Do you consider your child to be bi/multilingual? Yes ☐ No ☐

4. Have you ever contacted your GP, Speech and Language Therapy services or any other professional regarding concerns over your child's language? Yes ☐ No ☐

5. When did your child produce their first word?

Around their first birthday ☐
Around 18 months ☐
Around their second birthday ☐
At three years or later ☐

6. Does your child, or anyone in the family have a diagnosed language difficulty?

....

....

7. Has your child ever had difficulty with their hearing? Yes ☐ No ☐

If you would like to receive a summary of the results please provide your email address here:

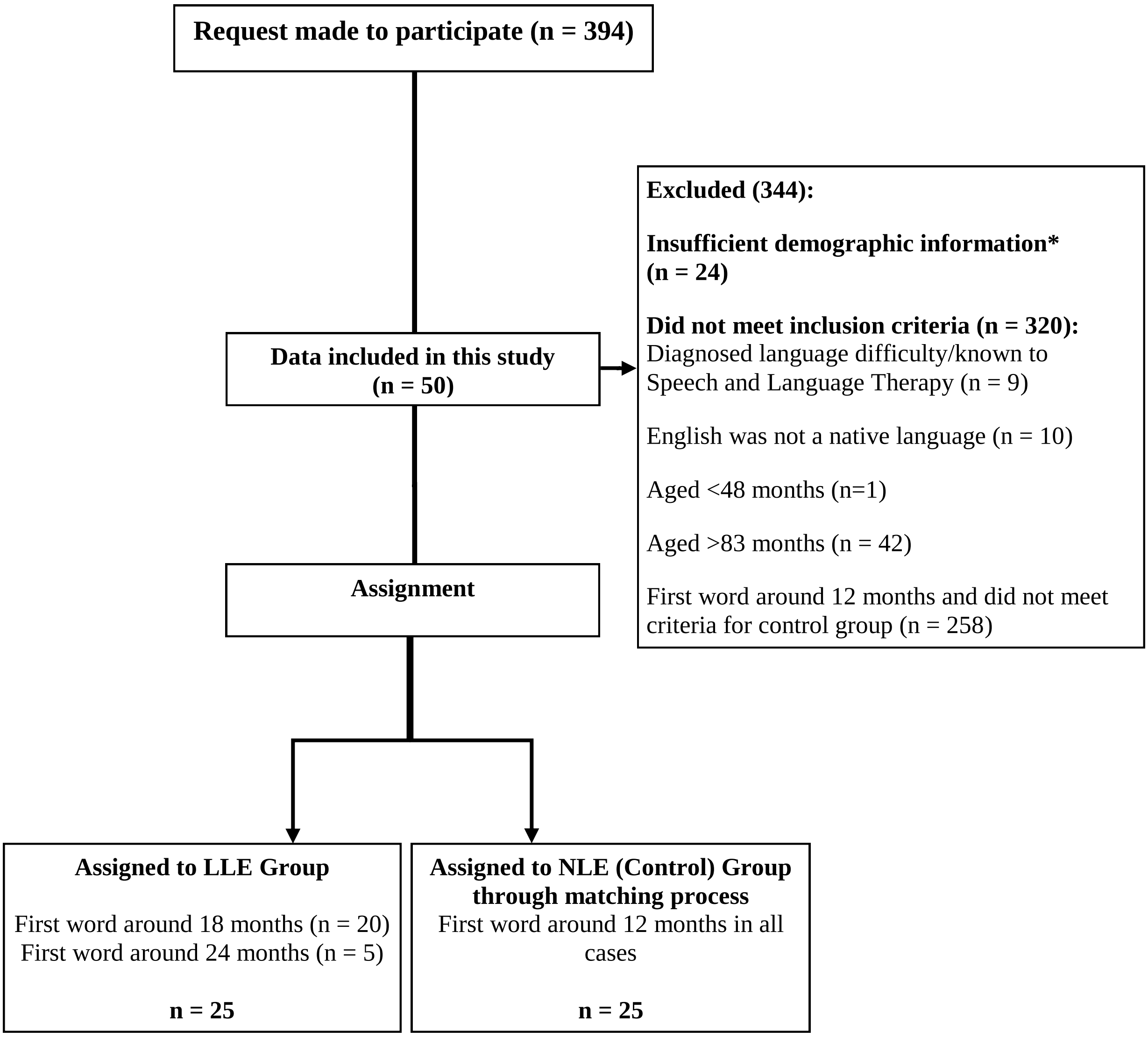
....

Thank you!

Note: Children's data was selected for inclusion in the LLE group if first word emerged around 18months or later; no language related difficulties were identified and English was a native language (mono-/multi-lingual).

Appendix B

Flow Chart Detailing Selection Process for Inclusion in the Study



Note. LLE = Late Language Emergence; NLE = Normal Language Emergence
*Questionnaires which omitted details relating to the inclusion criteria (age at first word; absence/presence of language difficulties; whether English was a native language) were considered incomplete.

Appendix C

Detailed participant's demographic information – arranged as matched pairs.

Participant	Age at 1st word (months)	Group	Age (months)	Gender	Multilingual	Family History
a2141	12	NLE	52	F	Yes	No
aa3205	18	LLE	49	F	Yes	No
b3026	12	NLE	52	F	No	No
bb3118	24	LLE	55	F	No	No
c4107	12	NLE	55	F	No	No
cc3189	18	LLE	55	F	No	No
d3180	12	NLE	58	F	Yes	No
dd2060	18	LLE	56	F	Yes	No
e3149	12	NLE	58	M	No	No
ee2139	24	LLE	58	M	No	No
f4155	12	NLE	58	M	No	No
ff2055	18	LLE	60	M	No	No
g4168	12	NLE	60	M	No	No
gg3052	24	LLE	60	M	No	No
h3109	12	NLE	65	F	Yes	No
hh3088	18	LLE	60	F	Yes	No
i2149	12	NLE	68	F	Yes	No
ii3202	18	LLE	61	F	Yes	No
j4133	12	NLE	75	F	Yes	No
jj4139	18	LLE	61	F	Yes	No
k4173	12	NLE	64	F	No	No
kk4029	18	LLE	64	F	No	No
l3173	12	NLE	70	F	No	Yes
ll4119	18	LLE	64	F	No	Yes
m2052	12	NLE	68	F	No	Yes
mm2072	18	LLE	67	F	No	Yes
n3103	12	NLE	67	F	No	No
nn4101	18	LLE	69	F	No	No
o4193	12	NLE	76	F	Yes	No
oo3028	18	LLE	70	F	Yes	No
p3194	12	NLE	75	F	No	No
pp3164	18	LLE	75	F	No	No
q4005	12	NLE	76	M	No	No
qq4131	18	LLE	75	M	No	No
r4043	12	NLE	76	M	No	No
rr2081	18	LLE	78	M	No	No
s3207	12	NLE	79	M	No	No
ss4104	18	LLE	79	M	No	No
t4034	12	NLE	80	F	No	No
tt4028	18	LLE	80	F	No	No
u3174	12	NLE	79	F	No	No
uu2036	24	LLE	82	F	No	No
v2069	12	NLE	77	F	No	No
vv3201	18	LLE	82	F	No	No
w3129	12	NLE	83	F	No	No
ww3021	18	LLE	83	F	No	No
x3199	12	NLE	64	M	Yes	No
xx3037	24	LLE	83	M	Yes	No
y4148	12	NLE	82	F	No	Yes
yy3062	18	LLE	83	F	Yes	Yes

Note. LLE = Late Language Emergence (first word at 18 or 24 months); NLE = Normal Language Emergence (first word at 12 months).

Appendix D

Complete data set. Participants were presented with either Alien 1 or Alien 2 data which contrasted audio and audiovisual presentation, ensuring all combinations were tested. All data was presented in three blocks, always starting with Block 1(2-3syllable non-words), then Block 2 (3-4 syllables) and finally Block 3 (4-5 syllables). Within each block non-words were presented in a pseudo-random order.

Stimulus	Data Set Alien 1	Data Set Alien 2	Block	Syllable no	AoA
meebip	A	AV	1	2	1
moobip	AV	A	1	2	1
bipoon	A	AV	1	2	1
bipeen	AV	A	1	2	1
geebong	A	AV	1	2	2
gooba	AV	A	1	2	2
tegong	A	AV	1	2	2
tegin	AV	A	1	2	2
shrosel	A	AV	1	2	3
throsel	AV	A	1	2	3
seslor	A	AV	1	2	3
zislom	AV	A	1	2	3
bopadomp	A	AV	1	3	1
podebomp	AV	A	1	3	1
pemeebind	A	AV	1	3	1
pomaybond	AV	A	1	3	1
keforka	A	AV	1	3	2
fekorfa	AV	A	1	3	2
tayfegor	A	AV	2	3	2
gootefor	AV	A	2	3	2
stothelo	A	AV	2	3	3
stoolero	AV	A	2	3	3
sheslaga	A	AV	2	3	3
sislorgi	AV	A	2	3	3
mopidoipe	A	AV	2	4	1
mubihoider	AV	A	2	4	1
woopeabi	A	AV	2	4	1
nubiapee	AV	A	2	4	1
keflechigo	A	AV	2	4	2
keflechorga	AV	A	2	4	2
takigafree	A	AV	2	4	2
jafitafree	AV	A	2	4	3
zoolerothay	A	AV	3	4	3
rolezootha	AV	A	3	4	2
thelereri	A	AV	3	4	3
reesareree	AV	A	3	4	3
dapawompibod	A	AV	3	5	1
padenindibid	AV	A	3	5	1
abrewdapopi	A	AV	3	5	1
abrombadopee	AV	A	3	5	1
trintecheviko	A	AV	3	5	2
tronkechafiki	AV	A	3	5	2
fegongtikitint	A	AV	3	5	2
tegongfetifint	AV	A	3	5	2
sharzirolethi	A	AV	3	5	3
sagilorathi	AV	A	3	5	3
ishilrethari	A	AV	3	5	3
ithilrasari	AV	A	3	5	3

Note. Presentation: A = Audio, AV = Audiovisual; Age of Acquisition (AOA): 1 = Early, 2 = Mid, 3 = Late