Compound and phrasal stress acquisition: When a greenhouse becomes different to a green house

Hannah Shilling

Supervised by: Andrea Krott

March 2010

Compound and phrasal stress acquisition: When a greenhouse becomes different to a green house

ABSTRACT

This study investigated the development of compound and phrasal stress in both comprehension and production, an area somewhat neglected by previous research of linguistic stress. 55 schoolchildren aged 4, 5 and 6 completed a picture-selection comprehension task and a picture-naming production task with minimal-pairs of compounds and phrases. The accuracy of picture selections in comprehension and of stress placement in productions revealed that 4-year-olds performed better on comprehension than production, indicating that production is refined later than comprehension, whilst 5- and 6-year-olds were better at comprehending phrases than compounds but better at producing compounds than phrases. Participant responses to compound and phrasal stimuli showed a phrasal-bias across all ages in the comprehension task regardless of stimulus type, possibly due to participant unfamiliarity with some compound word stimuli. In production 4-year-olds showed no response-bias for either stimulus type, whilst 5- and 6-year-olds demonstrated a bias towards compound stress placement. However, this only reached significance for compound stimuli but not phrasal, indicating that stress placement in participants' productions of phrases was more variable. These results demonstrate the complexities of compound and phrasal stress acquisition, and indicate that children's stress comprehension and production abilities are not adult-like by age 6. This suggests further improvement follows later in development.

KEY WORDS:	COMPOUND STRESS	PHRASAL STRESS	STRESS COMPREHENSION	STRESS PRODUCTION	STRESS ACQUISITION

Linguistic stress is a pervasive feature of verbal communication referring to the prominence placed on units of speech, which varies across words within a sentence and syllables within a word (Harley, 2008). Without stress, how would one know whether "it's hot" was a question or a statement? Or whether "record" meant 'a récord' (noun) or 'to recórd' (verb)? The former use of stress is intonation, used for emphasis within a sentence, whilst the latter is contrastive stress, whereby stress placement within a multisyllabic word directly affects meaning (Vogel & Raimy, 2002). Contrastive word stress can take another form: that which differentiates between compounds and phrases, i.e. distinguishing a gréenhouse (garden building) from a green hóuse (green-coloured house). It is this contrastive stress which the current study investigates.

In English, compound and phrasal stress are assigned by simple rules: the compound stress rule dictates that stress is on the first segment of a compound (Chomsky & Halle, 1968), whilst according to the nuclear stress rule phrasal stress is assigned to the rightmost phrase segment (Anttila, 2008; Chomsky & Halle, 1968). This difference in stress placement allows listeners to discriminate between compounds and phrases with identical constituents (Plag, 2006), such as the abovementioned "greenhouse" example.

Word stress is produced by manipulation of the acoustic factors fundamental frequency (F_0 - pitch), intensity (volume), and duration (word/syllable length) (Eilers, 1975; Fry, 1955; Kehoe, Stoel-Gammon & Buder, 1995; Kunter & Plag, 2007; Lieberman, 1960). Additionally, whilst unstressed syllables contain neutralised, reduced vowels, stressed syllables contain distinct full vowels (Pollock, Brammer & Hageman, 1993). Farnetani, Torsello & Cosi (1988) confirmed that adults produce compounds with higher F_0, intensity, and duration on the first syllable, whilst in disyllabic phrases these parameters are strongest on the second syllable.

Stress acquisition research suggests the ability to produce stress by manipulating these acoustic features improves developmentally with increasing prosodic and phonological control (Eilers, 1975; Patel & Brayton, 2009). Younger children primarily mark stress using duration, with reliable manipulation of F_0 and intensity following later (Patel & Brayton; Pollock et al., 1993), whilst reduction of unstressed syllables is the last skill to develop, possibly remaining unrefined until 12-years-old (Allen & Hawkins, 1980). There is considerable variability in research concerning the age at which children produce accurately-placed stress contrasts, with Pollock et al. finding that 4-year-olds show 99% accurate stress placement, whilst Patel & Brayton suggest that stress placement remains inaccurate and highly variable until at least age 7. Such discrepancy may result from the reliance of researchers such as Pollock et al. on children's repetitions of adult utterances, which may demonstrate increased accuracy due to imitation of adult stress placement. However, additional research by Clark, Gelman and Lane (1985) has shown that children as young as 3 can spontaneously produce novel compounds with accurate compound stress placement, although Wieman (1976) argues this reflects a tendency to stress new information, normally in the first segment of compounds, rather than knowledge of compounds.

In contrast to production it seems the ability to detect stress markers develops much earlier, with Jusczyk and Thompson (1978) and Spring and Dale (1977) using the high-amplitude sucking paradigm to demonstrate that one-month-old infants can discriminate between minimal-pairs of utterances differing only in stress placement. Furthermore, English infants appear able to distinguish the predominant stress pattern of their language (trochaic) by 7.5 months, shown by Kuhl (2004) using the head-turn paradigm, and are more likely to imitate stressed syllables than unstressed aged 28 months (Blasdell & Jensen, 1970). However, such research shows only infant ability to detect the acoustic correlates of stress, and fails to demonstrate when children can fully comprehend stress patterns, that is: at what age they use stress placement to determine meaning, hence the need to investigate children's ability to discriminate compounds and phrases.

To date, evidence of children's compound and phrasal stress comprehension appears limited to two studies: Atkinson-King (1973, cited in Allen & Hawkins, 1980 and Vogel & Raimy, 2002) and Vogel and Raimy (2002). Both studies employed a picture-selection task, in which children and adult controls were presented with one member of compound-phrase minimal-pairs as an auditory stimulus, following which they had to choose the correct interpretation from two pictures: one representing compound interpretation (e.g. hótdog; food type), the other representing phrasal (e.g. hot dóg; warm canine). Atkinson-King and Vogel & Raimy investigated children aged 4, 6, 8, 10 and 12, and 5, 7, 9 and 11, respectively, with both studies finding the youngest participants unable to use stress pattern to consistently select the correct interpretation. Comprehension accuracy increased with age until 11/12 years, at which point performance was adult-like: 100% in Atkinson-King's study but only 74% in Vogel and Raimy's. This lower optimum accuracy was attributed to the lack of explicit training in the latter study (which had been provided by Atkinson-King) and Vogel and Raimy's inclusion of novel compounds, which may have increased task difficulty.

Additionally, Vogel and Raimy's (2002) younger participants demonstrated a compound response-bias for real words regardless of stimulus type, indicating that learning to correctly interpret compound and phrasal stress requires overcoming this bias. Additionally, participants demonstrated a phrasal-bias when presented with novel compounds which was consistent across age, leading to two models of compound and phrasal stress interpretation (figure 1). According to the first model (figure 1a) when younger children interpret an auditory stimulus they disregard stress pattern, conducting a lexical search based only on segmental information; if they have a matching lexical entry the compound interpretation is selected, otherwise the constituent segments are searched for separately, leading to default phrasal interpretation. In contrast, older children and adults first distinguish the stress pattern of the stimulus and conduct their lexical search accordingly (figure 1b). However, despite distinguishing a compound stress pattern individuals may still select the default phrasal interpretation if they do not have a corresponding lexical entry, i.e. if a compound is unfamiliar.

Considering the aforementioned research, it seems clear that while linguistic stress has received much attention, the specific area of compound and phrasal stress acquisition has been somewhat neglected to date, with no study directly comparing both comprehension and production. Studies of compound production have focussed primarily on novel noun-noun compounds (e.g. Clark et al., 1985), not providing comparison between compound and phrasal stress placement. Although

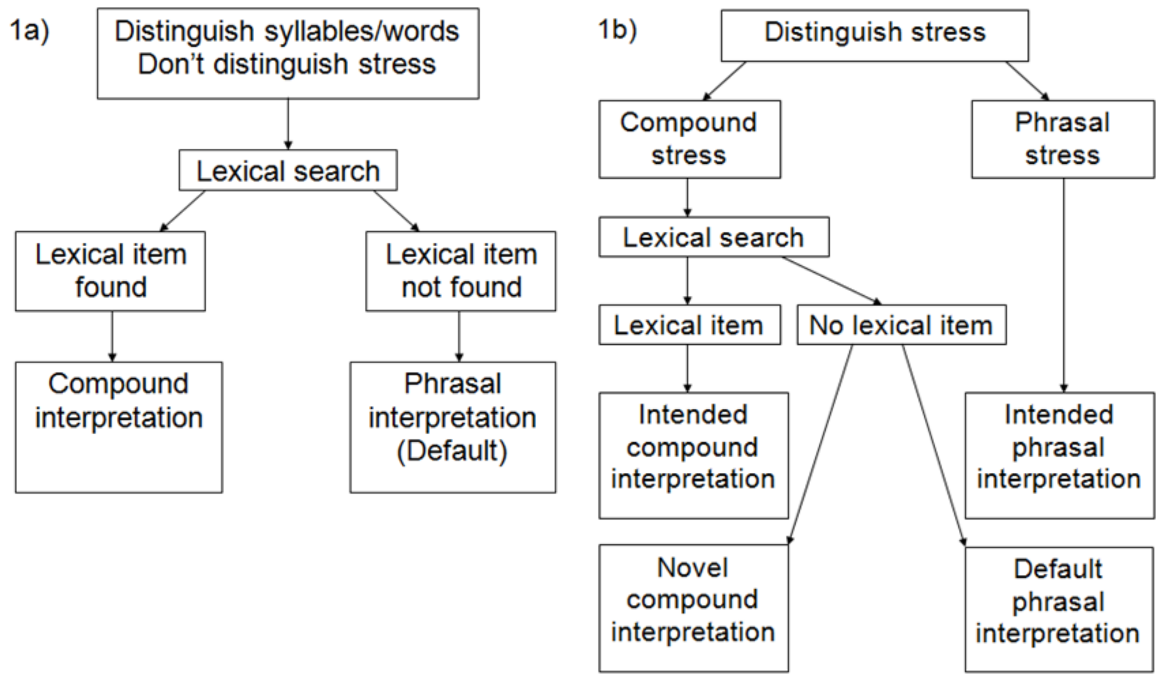


Figure 1: Models of stress distinction adapted from Vogel and Raimy (2002). 1a represents young children's failure to distinguish compound and phrasal stress. 1b represents older children and adult's ability to distinguish compound and phrasal stress

Atkinson-King (1973) and Vogel and Raimy (2002) provide more direct evidence of compound and phrasal stress comprehension, both studies neglected production and employed a forced-choice paradigm allowing only compound or phrasal interpretation, thus creating a high probability of correct picture selection regardless of whether participants understood the stress patterns, stimuli, or task. Additionally, each participant was presented with only one member of each minimal-pair, allowing no direct comparison between performance on two members of the same pair.

In order to address these concerns and investigate the matter of compound and phrasal stress acquisition further, the current study considers comprehension *and* production of *both* members of minimal-pairs of compounds and phrases. 4- to 6-year olds were tested using a picture-naming production task, and with a picture-selection comprehension task similar to that of Atkinson-King (1973) and Vogel and Raimy (2002), but with the inclusion of distracter items.

Considering previous findings, it was predicted that the youngest children would perform poorly on the comprehension task, demonstrating a compound-bias that decreases with age in accordance with Vogel and Raimy's (2002) proposed models. It was anticipated that performance would be poorer in production than comprehension, based on evidence that stress production is generally acquired and refined later than comprehension, but that performance on both tasks would improve with age. The stress patterns of responses given on the production task were analysed to determine whether a compound-bias similar to that previously demonstrated in comprehension is generalisable to production, or whether production develops differently.

Method

Design

This was a mixed-design study examining performance on two tasks (comprehension and production) with two stimulus types (compound and phrasal) across three different age groups (4-, 5- and 6-years-old).

Participants

The participants were 55 children from Reception to Year 2 classes at a primary school in Essex. All were native English speakers with no known disabilities, whose parents were issued information letters and returned consent forms. Participants were categorised according to age (Table 1).

Table 1. Participant information (ages shown as years; months)

Age group	Number of participants	Number of females	Number of males	Age range	Mean age
4	14	8	6	4;5 - 5;0	4;8
5	20	8	12	5;1 - 6;0	5;6
6	21	12	9	6;1 - 7;0	6;6

Materials

Both tasks used 16 minimal-pairs of compounds and phrases that differed only according to stress placement. All were adjective-noun phrases. Participants were tested with both members of each pair in both tasks.

Comprehension task. Auditory stimuli were recorded by the experimenter (a female native English speaker) in one of four neutral sentences (Appendix 1). Both members of each minimal-pair were recorded in the same sentence to control for context, and all stimuli were recorded at least 3 times, with the best stressed exemplar of each stimulus selected for experimental use. These were assessed by another linguistic researcher to ensure correct and natural-sounding stress placement.

For each auditory stimulus there were 4 corresponding picture cards, with simple coloured pictures on a plain white background. One picture represented the compound interpretation, another represented phrasal, and two were distracter items. These distracters represented the two constituent words of the stimulus, one representing the adjective, and the other representing the noun. For example, for the stimulus “green house” one distracter was something green, while another was a house (Appendix 2). Distracters were used to reduce the likelihood of participants selecting the correct picture by chance alone if they did not understand the stimuli or task.

Production task. In the production task children were presented with visual stimuli designed to elicit production of the desired phrases and compounds. In total 32 picture cards were used, a subset of those used in the comprehension task: 16

representing the compound interpretation of stimuli, and 16 representing phrasal (Appendix 3).

Procedure

Children were seen individually in a quiet room in their school, and each session lasted approximately 25-30 minutes. Children were seated adjacent to the experimenter at a table. Before beginning, the children were introduced to a teddy bear named “Fluffy” and asked if they would be willing to play a word and pictures game with Fluffy, to which all agreed. The order of comprehension and production tasks was counterbalanced across participants. In both tasks, compound and phrasal stimuli formed two separate blocks, and the order of these blocks was counterbalanced both within and across tasks, as was the order of words within these blocks. The order of words within the compound and phrasal blocks was matched so there was equal distance between the compound and phrasal member of each minimal-pair.

In the comprehension task participants were told that Fluffy would ask them some questions about pictures that would be placed in front of them, and they had to point to the picture Fluffy asked about. Participants were then presented with a set of four pictures on the table in front of them (the order of cards in the visual array was randomised across participants and stimuli) and the corresponding auditory stimulus was played through a CD player situated behind Fluffy. The participants’ picture selection was marked on a response sheet by the experimenter and later translated into computational data. This process was repeated until all stimuli had been completed. The first stimulus was preceded by a practice item which was neither recorded nor included in the analysis.

In the production task participants were told that Fluffy wanted to hear them say the name of some items. The experimenter held up one picture card at a time in front of the participant, coupled with the question “What’s this?”. Participants’ responses were recorded via a microphone on the table in front of them onto a Tascam digital voice recorder. Participants first completed a practice item, which was neither recorded nor analysed. For stimuli which participants could not name they were given a brief description of the item, followed by a phonetic clue to the word’s beginning, followed by clues to the word’s constituent segments. For items participants still could not name the experimenter pronounced the item’s name, asking the participant “Can you say that?”. Such items were marked as imitation responses.

These verbal responses were later analysed by the experimenter and a second listener, who rated each production for stress placement: compound, phrasal, or level. Agreement between listeners was 91.46%. An additional third listener rated those items the first two disagreed on; this lead to classification of a further 8.02% of all responses.

Results

Accuracy of responses. Following the process of rating stress placement in participants’ productions, those responses for which all three listeners disagreed on stress placement, or which two or more listeners rated as having level stress, were excluded from analysis (1.35% of responses), as were imitation items. Summary data indicated that the stimulus 'minibus' was unreliable in the comprehension task. Hence it was removed from both production and comprehension data prior to analysis. Unless otherwise stated all analyses employed an alpha level of .05.

To test the accuracy of participant responses a three-way split-plot ANOVA was conducted with age (3) as a between-participants factor and task type (2) and stimulus type (2) as within-participants factors, and number of correct responses as the dependent variable. This revealed significant main effects of age ($F(2,52) = 14.48$, $p<.001$, partial $\eta^2 = .36$) and task type ($F(1,52) = 5.18$, $p = .03$, partial $\eta^2 = .09$), with more correct responses with increasing age and slightly more correct responses in comprehension than in production, but not of stimulus type ($F(1,51) = 0.34$, $p = .56$, partial $\eta^2 = .01$). There were significant interactions between age and task type ($F(2,52) = 14.05$, $p<.001$, partial $\eta^2 = .35$), age and stimulus type ($F(2,52) = 6.75$, $p = .002$, partial $\eta^2 = .21$) and task type and stimulus type ($F(1,52) = 71.03$, $p<.001$, partial $\eta^2 = .58$). There was also a significant three-way interaction between age, task type and stimulus type ($F(2,52) = 8.96$, $p<.001$, partial $\eta^2 = .26$).

Table 2
T-tests comparing accuracy in response to compound and phrasal stimuli in comprehension and production tasks

Age Group	Paired Conditions		*t*	*df*	*Significance*
5	Comprehension: Compound	Production: Compound	-6.14*	19	<.001
5	Comprehension: Phrasal	Production: Phrasal	6.00*	19	<.001
5	Comprehension: Compound	Comprehension: Phrasal	-8.76*	19	<.001
5	Production: Compound	Production: Phrasal	4.19*	19	<.001
6	Comprehension: Compound	Production: Compound	-8.75*	20	<.001
6	Comprehension: Phrasal	Production: Phrasal	4.65*	20	<.001
6	Comprehension: Compound	Comprehension: Phrasal	-4.42*	20	<.001
6	Production: Compound	Production: Phrasal	6.85*	20	<.001

*Significant at the level p<.0125

Following these results subsequent within-participant stimulus type (2) x task type (2) ANOVAs were conducted for each age group to explore the interactions further. For 4-year-olds this revealed significant main effects of task type ($F(1,13) = 21.65$, $p<.001$, partial $\eta^2 = .63$) and stimulus type ($F(1,13) = 5.95$, $p = .030$, partial $\eta^2 = .31$), with more correct responses in the comprehension task than in the production task and more correct responses for phrases than compounds. There was no significant interaction between task type and stimulus type ($F(1,13) = 1.41$, $p = .256$, partial $\eta^2 = .10$). 5-year-olds showed no significant main effect of either stimulus type ($F(1,19) = 0.07$, $p = .79$, partial $\eta^2 = .004$) or task type ($F(1,19) = 0.03$, $p = .87$, partial $\eta^2 = .002$). For 6-year-olds there was a significant main effect of stimulus type ($F(1,20) =$

6.26, p = .021, partial η^2 = .24), with more correct responses for phrases than compounds, but not of task type (F(1,20) = 2.85, p = .11, partial η^2 = .13). There was also a highly significant interaction between task type and stimulus type for both 5- (F(1,19) = 42.22, p<.001, partial η^2 = .69) and 6-year-olds (F(1,20) = 71.91, p<.001, partial η^2 = .78).

To explore these interactions further, paired t-tests were run on the 5- and 6-year-old participants' data. These were all highly significant when an alpha level of .0125 was applied in accordance with the Bonferroni correction (Table 2).

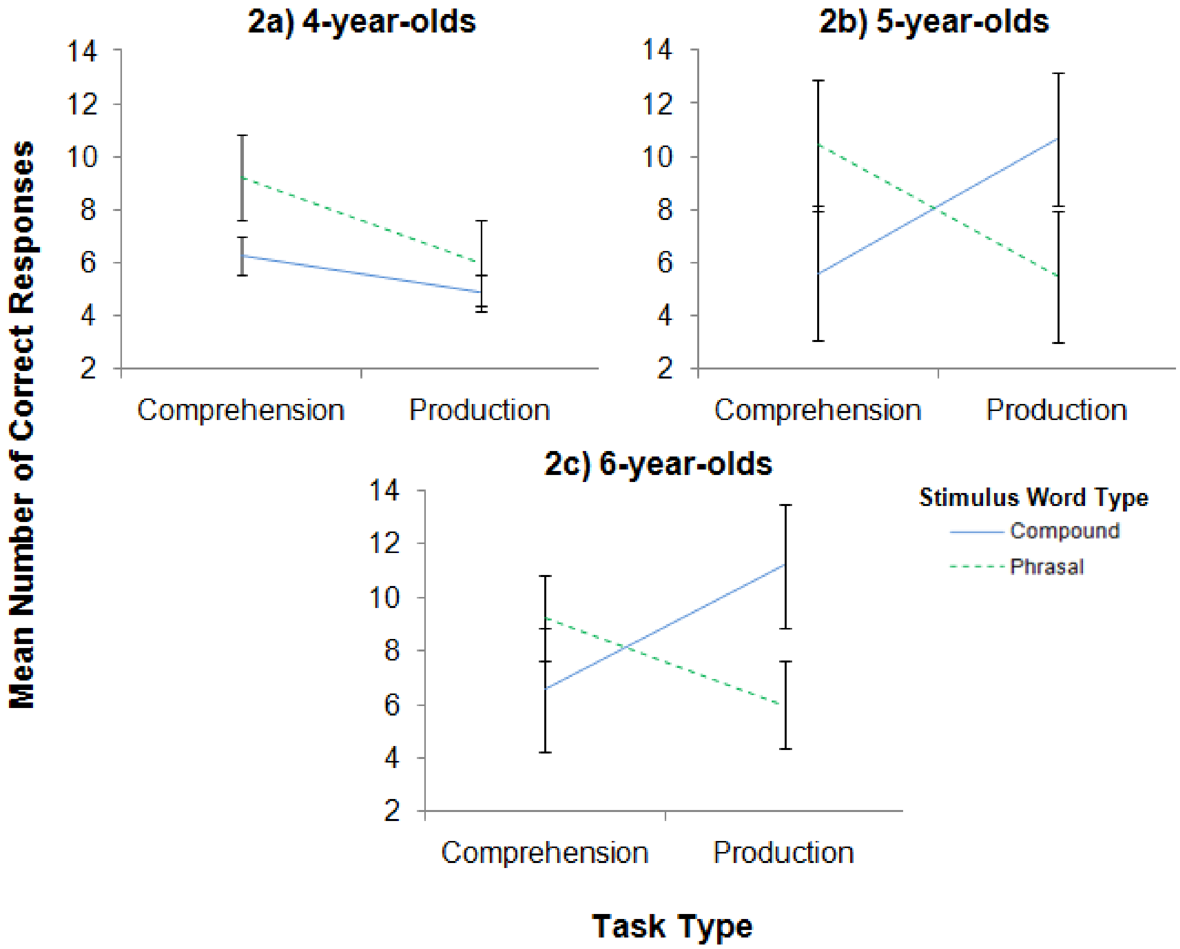


Figure 2: Accuracy of responses to compound and phrasal stimuli in comprehension and production tasks. 2a, 2b and 2c show 4-, 5-, and 6-year-olds' performance, respectively. Error bars represent standard error

Figure 2b and c shows that 5- and 6-year-olds performed better with phrasal stimuli in the comprehension task, and with compound stimuli in the production task. In contrast, figure 2a shows that 4-year-olds performed better with phrasal stimuli in both tasks. Figure 2 shows that performance with phrasal words in both tasks was consistent across age, as was performance with both word types in the comprehension task. However, it is clear that accuracy with compound words in the production task increased dramatically between the ages of 4 and 5 (figure 2a & b), with a further slight increase between the ages of 5 and 6 (figure 2b & c).

Meanwhile, comprehension of compounds remained consistently less than 7 accurate responses out of a possible 15, as did production of phrases (figure 2a-c).

Comprehension task. To test the number of each type of response (compound or phrasal) participants gave to each stimulus type in the comprehension task, a split-plot ANOVA was conducted with age (3) as a between-participants factor and stimulus type (2) and response type (2) as within- participants factors. This revealed significant main effects of response type ($F(1,52) = 70.75$, $p<.001$, partial $\eta^2 = .58$) and age ($F(2,52) = 7.23$, $p = .002$, partial $\eta^2 = .22$), with more phrasal responses than compound responses and fewer selections of distracter pictures with increasing age, and a significant interaction between stimulus type and response type ($F(1,52) = 44.77$, $p<.001$, partial $\eta^2 = .46$). All other results were non-significant (Table 3).

Table 3
Comprehension task: Non-significant results of age x stimulus type x response type ANOVA

Factor	*F*	*df*	*Significance*	Partial η^2
Stimulus type	0.729	1	0.40	0.01
Stimulus type x Age	1.914	1,52	0.16	0.07
Response type x Age	2.783	1,52	0.07	0.10
Stimulus type x Response type x Age	0.801	2,52	0.45	0.03

As there was no interaction involving age, all age groups were treated as one and paired t-tests were conducted to examine the stimulus type and response type interaction. These were all highly significant at the level $p<.0125$ (Table 4), with figure 3 demonstrating that all age groups were more likely to select a phrasal than compound interpretation regardless of stimulus type. Figure 3 also shows that participants were more likely to pick a compound response when presented with a compound stimulus than when presented with a phrasal stimulus, and more likely to pick a phrasal response when presented with a phrasal rather than compound stimulus.

Table 4
Comprehension task: T-tests comparing the number of compound and phrasal interpretations selected when presented with compound and phrasal stimuli. Condition names written as 'Stimulus type: Response type'

Paired Conditions		*t*	*df*	*Significance*
Compound: Compound	Compound: Phrasal	-5.03*	54	<.001
Phrasal: Compound	Phrasal: Phrasal	-10.15*	54	<.001
Compound: Compound	Phrasal: Compound	6.65*	54	<.001
Compound: Phrasal	Phrasal: Phrasal	-6.55*	54	<.001

*Significant at the level p<.0125

Production task. To test the number of each response type participants gave to each stimulus type in the production task, a split-plot age (3) x stimulus type (2) x response type (2) ANOVA was conducted. This revealed significant main effects of age (F(1,52) = 25.51, p<.001, partial η^2 = .50), with more appropriate responses with increasing age, stimulus type (F(1,52) = 92.95, p<.001, partial η^2 = .64), with slightly more response attempts to phrases than compounds, and response type (F(1,52) = 40.21, p<.001, partial η^2 = .44), with more compound than phrasal responses. There were significant interactions between age and stimulus type (F(2,52) = 35.45, p<.001, partial η^2 = .58), age and response type (F(2,52) = 4.98, p=.01, partial η^2 = .16), and stimulus type and response type (F(1,52) = 22.56, p<.001, partial η^2 = .30). There was no significant age x stimulus x response interaction (F(2,52) = 1.94, p = .15, partial η^2 = .07).

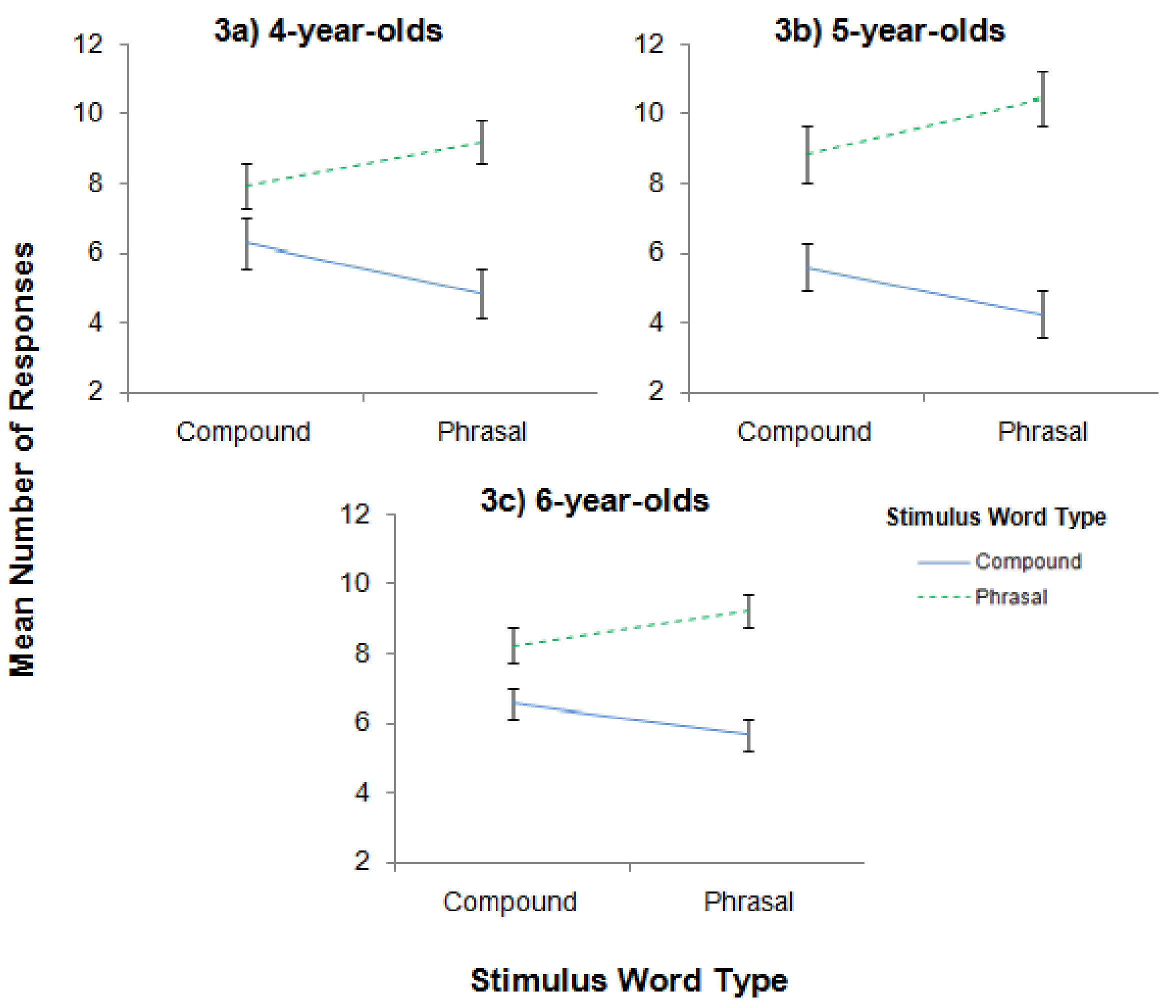


Figure 3: Comprehension task: Number of compound and phrasal picture responses participants selected when presented with compound and phrasal stimuli. 3a, 3b, and 3c show 4-, 5-, and 6-year-olds' responses, respectively. Error bars represent standard error

Following these results within-participants stimulus type (2) x response type (2) ANOVAs were conducted on each age group to further investigate the significant interactions. For all ages there was a significant main effect of stimulus type, with more responses to phrasal than compound stimuli [4-year-olds: (F(1,13) = 54.93, *p*<.001, partial η^2 = .81), 5-year-olds: (*F*(1,19) = 7.11, *p* = .02, partial η^2 = .27) and 6-year-olds (*F*(1,20) = 9.06, *p* = .01, partial η^2 = .31)]. At age 4 there was no significant main effect of response type (*F*(1,13) = 0.63, *p* = .44, partial η^2 = .05) and no significant stimulus type x response type interaction (*F*(1,13) = 0.81, *p* = .39, partial η^2 = .06). For both 5- and 6-year olds there was a significant main effect of response type, with more compound than phrasal responses [(*F*(1,19) = 20.94, *p*<.001, partial η^2 = .52) and (*F*(1,20) = 48.87, *p*<.001, partial η^2 = .71) respectively]. There was also a significant stimulus type x response type interaction for both age groups [aged 5: (*F*(1,19) = 27.19, *p*<.001, partial η^2 = .59), aged 6: (*F*(1,20) = 11.99, *p*=.002, partial η^2 = .38)].

Table 5

Production task: T-tests comparing the number of compound and phrasal responses produced when presented with compound and phrasal stimuli. Condition names in the form 'Stimulus type: Response type'

Age Group	**Paired Conditions**		*t*	*df*	*Significance*
5	Compound: Compound	Compound: Phrasal	6.39*	19	<.001
5	Phrasal: Compound	Phrasal: Phrasal	2.43	19	0.025
5	Compound: Compound	Phrasal: Compound	3.95*	19	0.001
5	Compound: Phrasal	Phrasal: Phrasal	-6.20*	19	<.001
6	Compound: Compound	Compound: Phrasal	10.54*	20	<.001
6	Phrasal: Compound	Phrasal: Phrasal	2.20	20	0.040
6	Compound: Compound	Phrasal: Compound	2.95*	20	0.008
6	Compound: Phrasal	Phrasal: Phrasal	-3.96*	20	0.001

*Significant at the level p<.0125

To explore these interactions further, paired t-tests were run on 5- and 6-year-olds' responses using an alpha level of .0125 (Table 5). Figure 4b & c shows both age groups were significantly more likely to produce a compound response to a compound rather than phrasal stimulus, and to produce a phrasal response to a phrasal rather than compound stimulus. Additionally, 5- and 6-year olds were significantly more likely to produce a compound response than phrasal when presented with a compound stimulus. Both groups were also more likely to produce a compound response than phrasal even when the stimulus was phrasal, indicating a compound-bias, but this trend did not reach significance, meaning that 5- and 6-year-olds' productions in response to phrasal stimuli were more variable than their responses to compound stimuli. In contrast, figure 4a shows that 4-year-olds' productions were variable in response to both stimulus types, with participants not showing a preference for either response type when presented with either stimulus type.

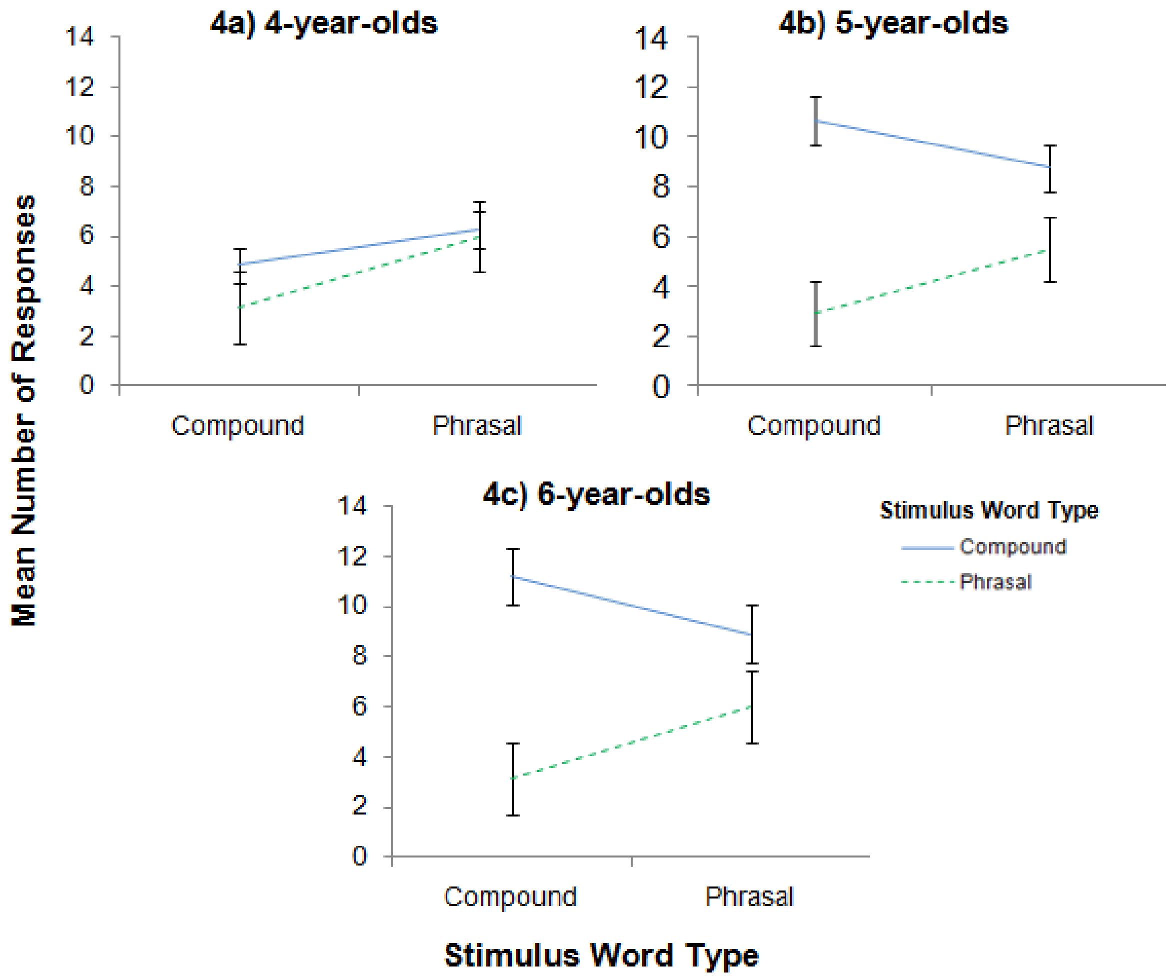


Figure 4: Production task: Number of compound and phrasal verbal responses produced when presented with compound and phrasal stimuli. 4a, 4b, and 4c show 4-, 5-, and 6-year-olds' responses, respectively. Error bars represent standard error

Discussion

The current study aimed to investigate comprehension and production of minimal-pairs of compounds and phrases in 4- to 6-year-olds to determine at what age children develop the ability to differentiate between compounds and phrases based on stress pattern alone.

In agreement with our hypothesis, overall performance was more accurate in the comprehension task than production. However, it is important to note that this trend was entirely due to 4-year-olds' highly inaccurate performance with compounds in the production task, as overall 5- and 6-year-olds performed equally well on both tasks, demonstrating higher accuracy with compounds than phrases in the production task, but higher accuracy with phrases than compounds in the comprehension task. This poorer performance of 4-year-olds compared to 5- and 6-year-olds in production but not comprehension supports the suggestion of previous research that stress production is refined later than comprehension (Allen & Hawkins, 1980). This may be because knowledge of the relation between prosody and semantics develops considerably between the ages of 4 and 5, allowing children to purposely produce more accurately-placed stress (Cutler & Swinney, 1987). The low accuracy of 4-year-old participants' productions could also be because these children were in their first year of school, at which point they may be more concerned with vocabulary expansion than the specificities of production. As Pikulski & Templeton (2004) noted, upon first beginning school there is a large focus on expanding children's written vocabulary to match their verbal abilities, hence accurate stress placement may be relatively unimportant to children at this age.

It was additionally predicted that performance accuracy in both tasks would increase with age, as stress acquisition improves developmentally (Patel & Brayton, 2009; Pollock et al., 1993; Atkinson-King, 1973; Vogel & Raimy, 2002), and does not reach optimum level of accuracy by age 4 (Atkinson-King; Patel & Brayton; Vogel & Raimy) hence would be expected to improve within the investigated age range. The previously mentioned improvement in production accuracy, particularly between the ages of 4 and 5, complies with this prediction, supporting previous findings of improved stress placement accuracy with age (Kehoe et al., 1995; Patel & Brayton) and providing additional evidence specific to compounds and phrases. That is, there is significant improvement in production of compounds between the ages of 4 and 5, whilst the production of phrases remains largely inaccurate aged 4-6, suggesting improvement in phrase production follows later.

On the other hand there was no significant improvement in comprehension accuracy between the ages of 4 and 6, with accuracy already relatively high by age 4, particularly for phrases. Unlike Atkinson-King (1973) and Vogel and Raimy's (2002) findings this suggests that 4-year-olds have already achieved a certain level of competence when comprehending compounds and phrases. This implies that studies showing infant (Spring & Dale, 1977) and young children's (Blasdell & Jensen, 1970) distinction of stress patterns may indeed represent the beginnings of stress comprehension rather than only the ability to distinguish acoustic features. However, none of the current age groups demonstrated complete accuracy in the comprehension task, with less than 50% accuracy when comprehending compounds. Compared with the 74% and 100% comprehension accuracy shown in Vogel and Raimy and Atkinson-King's studies, these results suggest there is further improvement in comprehension of compound and phrasal stress later in development, despite stability across the investigated ages. Alternatively, participants may have been performing at the optimum level for stimuli that they knew, but were possibly unfamiliar with some stimuli. The implications of this are explained below.

The phrasal-bias demonstrated by the current participants in the comprehension task directly contradicts the compound-bias shown by Vogel and Raimy's (2002) participants and may be due to participants misunderstanding the stress patterns of stimuli or failing to bind the constituent segments of compounds prior to conducting their lexical search. However, this interpretation appears unlikely as compounds have a high level of internal cohesion, particularly in comparison to phrases (Farnetani et al., 1988). Alternatively it may be that the stimuli used in this investigation overestimated the vocabulary of the participants, i.e. some of the compound stimuli used were unfamiliar to participants. This latter explanation appears more plausible, and can be accommodated by Vogel and Raimy's proposed model of older children's/adults' stress distinction (refer back to figure 1b). Assuming that the current participants were unfamiliar with some compound stimuli, they may have correctly discriminated a compound stress pattern, yet still reached a phrasal interpretation by default due to their lack of matching lexical entry. The current phrasal-bias resembles that shown by Vogel and Raimy's participants in response to novel stimuli, further implying that the current phrasal-bias resulted from participants' unfamiliarity with some compound stimuli.

The current study also aimed to consider whether a bias would be evident in 4-6-year-olds' productions of compounds and phrases. The current data is split by age, with 4-year-olds not demonstrating a bias for either response type when presented with either stimulus type, illustrating great variability in the stress placement of their productions. This further supports the aforementioned inference that stress placement may be rather unimportant for children at this age. In contrast, 5- and 6-year olds demonstrated a compound-bias in response to compound stimuli, but higher variability when presented with phrasal stimuli. This finding is not necessarily surprising, as although there is a tendency for phrases to be produced with stronger stress on the last syllable this is not an invariant pattern (Hayes, 1995), and often both constituents of a phrase are stressed (Farnetani et al., 1988). Therefore it is not necessarily wrong for phrases to receive first-constituent stress.

Overall, the current study has produced several findings to contribute to the relatively neglected area of compound vs. phrasal stress acquisition, whilst overcoming the limitations of previous studies by inclusion of distracter items and direct comparison of comprehension and production. However, there are limitations to this study that future research should seek to improve. Firstly, the abovementioned assumption that some of the current results were due to participant unfamiliarity with compound test stimuli cannot be confirmed, as no information was gathered concerning participants' vocabulary. To overcome this, one would need to ascertain each individual's familiarity with word stimuli prior to testing, for example using parental questionnaires of participants' lexical knowledge, with the inclusion of dummy items and firm instructions not to purposely teach the children those items which they may not already know. Removal of unfamiliar words prior to analyses would demonstrate whether the same phrasal-bias is evident in comprehension when unfamiliarity is not an issue.

Additionally, although auditory stimuli for the comprehension task were carefully selected to ensure correct compound and phrasal stress, participants may possibly have shown a phrasal comprehension bias as they perceived compound stress as intended to differentiate between phrasal and distracter items. For example, participants may have perceived the stress on "gréen" in "gréenhouse" as intended to differentiate between the phrasal visual stimuli of a *green* house and the second constituent distracter item, a *white* house (see appendix 2i). To test whether this had occurred it would be beneficial to run trials in which the distracter items were unrelated to phrase constituents to determine whether this phrasal-bias persists, and hence cannot be attributed to distracter/phrase comparison.

A further limitation is the lack of adult control group, which disallows one from making direct comparisons between 4- to 6-year-old performance and the optimum level of accuracy shown by adults. Therefore, the current study can only speculate whether the performance shown by 4- to 6-year-olds is adult-like and about further improvement in accuracy after age 6. While adult performance in previous comprehension studies suggests accuracy increases further, previous levels of optimum performance cannot truly be applied to this study, as adult accuracy appears to vary dependent on methodological differences [i.e. Atkinson-King (1973) vs. Vogel & Raimy (2002)]. The lack of existing production research and previously reported variability in even adult phrase productions (Farnetani et al., 1988; Hayes, 1995) makes this issue of an adult control group especially important in the production task.

Therefore, further research concerning compound and phrasal stress acquisition should aim to address the above limitations, possibly also extending the age range considered to allow comparison of comprehension and production throughout later development. Attempting to replicate the current findings in a more natural setting may also be beneficial, for example in a play session where auditory stimuli are verbalised by an experimenter in real-time and visual stimuli take the form of toys. This would enhance ecological validity and reduce any testing effects.

In conclusion, although not without limitation the current study provides evidence to enhance knowledge of compound and phrasal stress acquisition. The current findings both support and question previous evidence, suggesting that previous findings of chance accuracy in comprehension aged 4 may have underestimated children's comprehension abilities, whilst the variability of 4-year-olds' productions demonstrates that the general trend of later refinement of production than comprehension of stress is also applicable to compounds and phrases. Although the phrasal-bias demonstrated by participants contradicts previous findings, this can be sufficiently accommodated by Vogel and Raimy's (2002) models of stress distinction, hence supporting their proposed model and highlighting a methodological flaw which may have made compound stress comprehension abilities aged 4-6 appear weaker than they actually are. The evidence of heightened variability in stress placement within phrases compared to compounds provided here contributes to our understanding of phrasal stress, and may warrant further exploration to determine at what age children become more patterned and less variable in their productions of phrases. Thus, while the current study's consideration of both comprehension and production provides valuable addition to knowledge of compound and phrasal stress acquisition, it also raises questions about further development of compound and phrasal stress beyond the age of six. Future researchers should continue investigation within this area to provide more definitive answers.

References

Allen, G., & Hawkins, S. (1980). Phonological rhythm: Definition and development. In G. Yeni-Komshian, J. Kavanagh & C. Ferguson (Eds.), *Child Phonology: Volume 1. Production* (pp. 227-256). New York, NY: Academic Press.

Anttila, A. (2008). Phonological constraints of constituent ordering. In C. B. Chang & H. J. Hayne (Eds.) *Proceedings of the 26th West Coast Conference on Formal Linguistics.* Somerville, MA: Cascadilla Proceedings Project, pp. 51-59.

Atkinson-King, K. (1973). Children's acquisition of phonological stress contrasts. *Unpublished Doctoral Dissertation:* University of California, Los Angeles.

Blasdell, R., & Jensen, D. (1970). Stress and word position as determinants of imitation in first-language learners. *Journal of Speech and Hearing Research, 12,* pp. 193-202.

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

Clark, E. V., Gelman, S. A., & Lane, N. M. (1985). Compound nouns and category structure in young children. *Child Development, 56,* pp. 84-94.

Cutler, A., & Swinney, D. A. (1987). Prosody and the development of comprehension. *Journal of Child Language, 14,* pp.145-167.

Eilers, R. E. (1975). Suprasegmental and grammatical control over telegraphic speech in young children. *Journal of Psycholinguistic Research, 4,* pp. 227-239.

Farnetani, E., Torsello, C. T., & Cosi, P. (1988). English compound versus non-compound noun phrases in discourse: An acoustic and perceptual study. *Language and Speech, 31*, pp. 157-180.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America, 27,* pp.765-768.

Harley, T. A. (2008). *The Psychology of Language: From Data to Theory* (3rd Edition).Hove, UK: Psychology Press.

Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies.* Chicago, IL: The University of Chicago Press.

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

Kehoe, M., Stoel-Gammon, C., & Buder, E. H. (1995). Acoustic correlates of stress in young children's speech. *Journal of Speech and Hearing Research, 38,* pp. 338-350.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5,* pp. 831-843.

Kunter, G., & Plag, I. (2007). What is compound stress? In J. Trouvain & W. J. Barry (Eds.) *Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken,* pp. 1005-1008.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America, 32,* pp. 451-454.

Patel, R., & Brayton, J. T. (2009). Identifying prosodic contrasts in utterances produced by 4-, 7-, and 11-year-old children. *Journal of Speech, Language and Hearing Research, 52,* pp.790-801.

Pikulski, J. J., & Templeton, S. (2004). *Teaching and developing vocabulary: Key to long-term reading success.* Boston, MA: Houghton Mifflin Company.

Plag, I. (2006). The variability of compound stress in English: Structural, semantic and analogical factors. *English Language and Linguistics, 10,* pp.143-172.

Pollock, K. E., Brammer, D. M., & Hageman, C. F. (1993). An acoustic analysis of young children's productions of word stress. *Journal of Phonetics, 21*, pp.183-203.

Spring, D. R., & Dale, P. S. (1977). Discrimination of linguistic stress in early infancy. *Journal of Speech and Hearing Research, 20,* pp.224- 232.

Vogel, I., & Raimy, E. (2002). The acquisition of compound vs. phrasal stress: The role of prosodic constituents. *Journal of Child Language, 29,* pp.225-250.

Wieman, L. A. (1976). Stress patterns in early child language. *Journal of Child Language, 3,* pp.283-286.

Appendices

Appendix 1: Comprehension task auditory stimuli – script of recordings

Compound stimuli	Phrasal stimuli
Which one is the Bigbird?	Which one is the big bird?
Can you find the blackberry?	Can you find the black berry?
Can you point to the bluebell?	Can you point to the blue bell?
Where is the blueberry?	Where is the blue berry?
Which one is the bluebird?	Which one is the blue bird?
Can you point to the bluebottle?	Can you point to the blue bottle?
Can you find the goldfish?	Can you find the gold fish?
Where is the greenfly?	Where is the green fly?
Can you point to the greenhouse?	Can you point to the green house?
Can you find the highchair?	Can you find the high chair?
Which one is the hotdog?	Which one is the hot dog?
Where is the lighthouse?	Where is the light house?
Can you point to the minibus?	Can you point to the mini bus?
Which one is the redhead?	Which one is the red head?
Can you find the shortbread?	Can you find the short bread?
Where is the whiteboard?	Where is the white board?

Appendix 2: Comprehension task picture stimuli – Shown in the order Compound interpretation; Phrasal interpretation, Distracter 1; Distracter 2

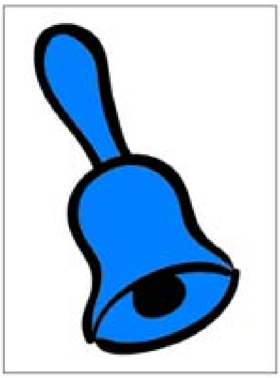
a) Bigbird/Big bird

b) Blackberry/Black berry

c) Bluebell/Blue bell

d) Blueberry/Blue berry

e) Bluebird/Blue bird




f) Bluebottle/Blue bottle




g) Goldfish/Gold fish




h) Greenfly/Green fly

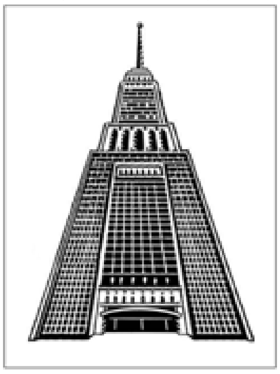
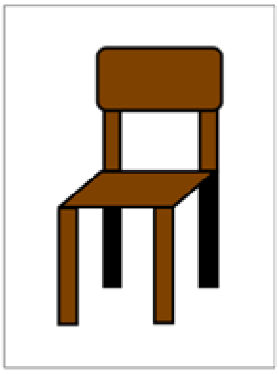



i) Greenhouse/Green house

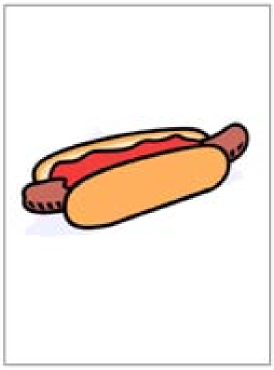


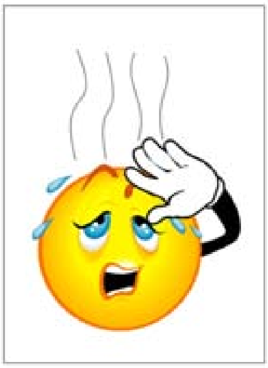

j) Highchair/High chair



k) Hotdog/Hot dog






l) Lighthouse/Light house




m) Minibus/Mini bus




n) Redhead/Read head




o) Shortbread/Short bread




p) Whiteboard/White board




Appendix 3: Production task picture stimuli

a) Compound stimuli

Bigbird

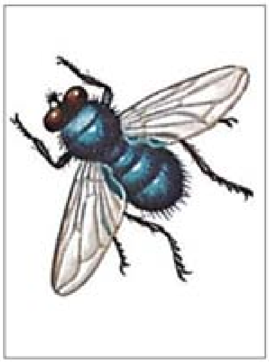
Blackberry

Bluebell

Blueberry

Bluebird

Bluebottle



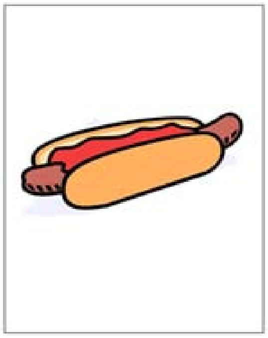
Goldfish

Greenfly

Greenhouse

Highchair

Hotdog



Lighthouse

Minibus

Redhead

Shortbread

Whitebread

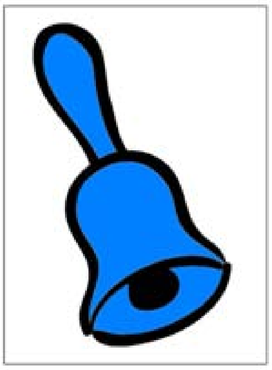



b) Phrasal stimuli

Big bird

Black berry

Blue bell



Blue berry

Blue bird

Blue bottle

Gold fish

Green fly

Green house

High chair

Hot dog

Light house

Mini bus

Red head

Short bread

White bread