PHONOLOGICAL ‘WILDNESS’ IN EARLY LANGUAGE DEVELOPMENT: EXPLORING THE ROLE OF ONOMATOPOEIA

CATHERINE E. LAING

University of York

Abstract

This study uses eye-tracking to single out the role of ‘wild’ onomatopoeia in language development, as described by Rhodes (1994). Wildness–whereby extra-phonetic features are used in order to reproduce non-human sounds–is thought here to facilitate infants’ understanding of onomatopoeic word forms, providing a salient cue for segmentation and understanding in the input. Infants heard onomatopoeic forms produced in familiar and unfamiliar languages, presented in a phonologically ‘wild’ (W) or ‘tame’ (T) manner. W forms in both familiar and unfamiliar languages were hypothesised to elicit longer looking times than T forms in both familiar and unfamiliar languages. Results reflect the role that onomatopoeia play in early language development: wildness was not found to be a factor in infants’ understanding of word forms, while reduplication and production knowledge of specific stimuli generated consistent responses across participants.

1. Introduction

Onomatopoeia appear amongst the early words of infants acquiring many languages, yet it is not uncommon for studies of child language development to disregard these word forms when analysing infant speech (e.g., Behrens, 2006, Genesee et al., 2008). While onomatopoeia could be considered as marginal to the adult language, they often constitute a considerable portion of an infant’s first word forms (Kauschke & Klann-Delius, 2007, Menn & Vihman, 2011 (see appendix)), and a focus on this early vocabulary may explain some of the developments that follow as an infant’s lexicon progresses towards the adult model.

Onomatopoeia are derivative of sound symbolism, which is a fully integrated feature of many languages, including Korean and Japanese (Ivanova, 2006). Sound symbolism, or ‘mimetics’, draws from the phonetic properties of a word to represent the synesthetic features of the object or state that it describes (Rhodes, 1994), resulting in a highly expressive parallel lexicon which is fully established as part of the language (e.g., Japanese ‘pika’ *a flash of light*, ‘goro’ *a heavy object rolling* (Kita, 1997)). Onomatopoeic word forms differ in that they constitute phonetic imitations of sounds in the environment, produced within the limits of the vocal tract. Unlike mimetics, onomatopoeia do not express physical features through the word’s phonetic or phonological properties, but rather they serve to phonetically reproduce non-human sounds (e.g., ‘thud’, ‘vroom’).

The use of mimetics in language development has been well-documented in the literature, found to facilitate the learning of Japanese novel verbs amongst Japanese and English-speaking adults and infants (Imai *et al*., 2008, Kantarzis *et al*., 2011). Mimetic forms appear at the very onset of Japanese infants’ word production, where they are used with a high level of accuracy (Tsujimura, 2005), and become increasingly complex over time as the infant acquires a full lexicon of both mimetic and non-mimetic words (Iwasaki *et al*., 2007). This

Proceedings of the first Postgraduate and Academic Researchers in Linguistics at York (**PARLAY 2013**) conference
ISSN 1758-0315 Issue 1

evidence towards a role for mimetics in early word learning suggests that onomatopoeic words may similarly support the learning of new word forms in the early output.

Whether the infant is acquiring a language that is rich in sound symbolism, such as Japanese, or a language which contains little (if any) sound symbolism, such as English, it appears that sound symbolic words–both mimetic and onomatopoeic–could be perceptually salient to infants acquiring their first word forms. Rhodes (1994) describes a model of 'wild' and 'tame' onomatopoeia, which explains the extent of phonetic imitation found in the features of the word form. 'Tame' forms are produced within the phonetic norms of the ambient language, adhering to normalised phonological structures that are familiar to the speaker, while 'wild' forms make use of the vocal tract's full capacity in order to approximate as closely as possible to the sound that the speaker is imitating. Wild forms draw upon vocal gestures that are not ordinarily used in the adult language, raising the question as to whether it is precisely these phonetic 'special-effects' that render onomatopoeic forms more salient in the speech stream, thus facilitating perception, memory, and eventually production of infants' earliest word forms. Wild onomatopoeia do not correspond to the typical segments and syllable-structures of the adult language, and may provide a perceptual attractor for infants as they attend to the speech stream.

This study uses eye-tracking to address infants' perception of wildness in onomatopoeic forms, which is presumed to provide a highly salient linguistic 'hook' in the input, facilitating lexical memory and the formation of word representations in language development. Wild features are assumed here to provide prosodic cues in the input, while being easier to recall than the typical native language phonology, to which tame forms adhere. It is hypothesised that infants will respond most systematically to the wild forms, thought to be easily recognisable due to their idiosyncratic 'special effect' features. This would indicate that infants respond most readily to the linguistically atypical features of the speech stream when acquiring language, underlining those features which are essential to the earliest stages of language development, but which do not necessarily match the words or phonemes that will eventually form the adult output.

2. Method

2.1. Participants

Nineteen Swedish infants (10 male, 9 female) between the ages of 14 and 16 months were tested (mean age 461.5 days). Infants were all full-term, and acquiring Swedish as their first language. A further five infants participated in the experiment but were excluded from the analysis due to fussiness during the eye-tracking procedure (4) or experimenter error (1).

2.2. Stimuli

Six onomatopoeic words (OWs)–all animal sounds[1]–which consistently appeared on English, Swedish, German and French adaptations of the MacArthur-Bates Communicative Development Inventory (CDI, Fenson *et al*., 1994) were selected for use in the experiment to ensure that participants were likely to have had prior experience of the stimuli. Two different photographic images of each of the corresponding animals were selected: the animals were all stood facing in the same direction, looking towards the infant from the right hand side, and

---

[1] The OW equivalents of COW, SHEEP, DOG, CAT, DUCK and ROOSTER were used in the experiment.

presented on a grey background. OWs were recorded in Swedish (the familiar language, L_F), and in languages unfamiliar to the infants (L_U)–Chinese, Arabic and Urdu.

Audio stimuli were recorded by native speakers, all female postgraduate students in the Linguistics departments at the universities of York or Stockholm. Each student was first asked to produce the OW as they would produce it when speaking to a toddler, as if imitating the animal in question ('wild' W). The students were then asked to produce the words with no added prosodic features, keeping to the natural phonology and stress pattern of their native language ('tame' T). Each word was produced once in each recording, adhering to the conventional full form of the word; words which would normally undergo reduplication were reduplicated (e.g. *quack quack*), while those which the speakers deemed as having no reduplication in production were recorded without reduplication (e.g. *cock-a-doodle-doo*).

Four adults, none of them speakers of any of the L_U languages, were then tested on their recognition of the W_U and T_U stimuli prior to the analysis. Only one of the 24 stimuli was found to be unrecognisable by all of the adults, which was removed from the analysis. Of the seven stimuli that were judged incorrectly by at least one of the adults, all but one were produced in a T manner. These results confirmed the suitability of the stimuli used in the infant experiment, as well as supporting the hypothesis that wildness facilitates word recognition.

2.3. Procedure

The experiment was controlled using E-Prime, with the visual stimuli presented using a 17" Tobii Studio 1750 eye-tracking monitor. Caregivers held the infant on their laps in a chair placed in front of the screen, and a five-point infant calibration was taken for each participant before the experiment began. The experimental procedure lasted approximately four minutes, during which time the caregiver was asked to wear headphones playing music from a Swedish radio station.

The experiment consisted of a salience phase and a test phase: during the salience phase pairs of images were displayed on the screen for 4000ms, before a centralising image of a baby appeared in the middle of the screen which served to 'reset' the infants' eye-gaze prior to the test phase. The image disappeared automatically upon fixation (or after 4000ms if the infant did not fixate), and the OW was heard through speakers on either side of the screen immediately after offset of the fixation image; the test phase lasted for 3000ms. After the experiment infants were rewarded with a certificate and parents were asked to complete a Swedish CDI questionnaire.

Each infant heard a total of 24 OWs: each of the four conditions (W_F, W_U, T_F and T_U) for each of the six animals, with a distribution of all three unfamiliar languages across the stimuli. The order of data output and the target's location on screen was randomised using E-Prime. Selection of the distractor image was partially randomised in E-prime according to the size of animal in the target image: to ensure against confusion between the images (e.g. sheep and dog, duck and rooster), animals were grouped into two categories–'small' and 'large'–and for each trial the distractor image was chosen from the opposite category to avoid ambiguity.

3. *Results*

Fixations during test phase were analysed from a window of 300-1800ms after onset of the stimulus. The proportion of looking towards the target was calculated for each trial as a percentage of the total fixation time for both target and distractor, and a mean looking time was calculated for each infant in each condition.

3.1. *Wildness and Familiarity*

A two-way repeated measures ANOVA was carried out with two factors: wildness (W vs. T) and familiarity (L_F vs. L_U), with proportion of looking towards the target image as the dependant variable (n = 19). This revealed no significant effect for wildness (F(1, 18) = 3.428, p = .081) or familiarity (F(1, 18) = .486, p = .495). The interaction was not significant either (F(1, 18) = 1.617, p = .220), and as is evident in Figure 1, results were around chance (0.50) for all conditions.

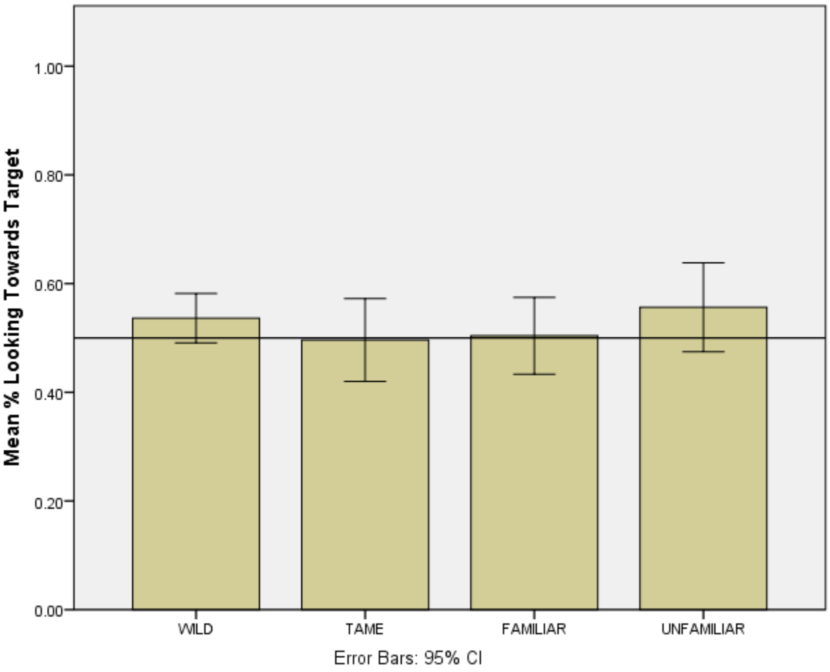

Figure 1: Results for all infants across conditions

These results raise the question of the interpretation of 'familiarity' in this experiment: were the stimuli really familiar to the infants, and if so, how familiar? In order to address this question, it is necessary to investigate the infants' knowledge of the individual OW forms used in the experiment. Results from the CDI questionnaires were used to determine infants' knowledge of individual word forms, both in terms of the OW, and the conventional word (CW) equivalent (for example, *woof* versus *dog*).

3.2. *Knowledge of Stimuli*

Breaking down the findings in this manner made it possible to explore the results in more depth. Infants were given two knowledge scores for each of the six target words–one for the OW and one for the CW–based on whether or not they were able to produce the word form. Results were then separated into knowledge groups in accordance with these scores. This approach was based on the assumption that being able to produce the CW would only strengthen an infant’s representation of its OW counterpart, while also suggesting that an infant is more advanced in his language development when he has started to produce CWs. The scoring conventions are presented in Table 1.

	Produces word	Doesn’t produce word[2]
OW	1	0
CW	1	0

Table 1: Knowledge Scoring for OWs and CWs

Scores were allocated for individual stimuli, meaning that an infant may be in OW1 and CW0 group for DOG if he is able to produce *woof* but not *dog*. The average looking time towards target for infants in each knowledge group was calculated.

3.2.1. Knowledge 0

The CW0 group (n = 17) was analysed with a two-way repeated measures ANOVA using the same two factors, and a significant interaction between wildness and familiarity was found (F(1, 16) = 8.557, p = .01). As shown in Figure 2, infants who were not able to produce the CW looked longer than chance at L_U words only. This is illustrated more clearly in Figure 3, where familiarity can be seen to interact with wildness.

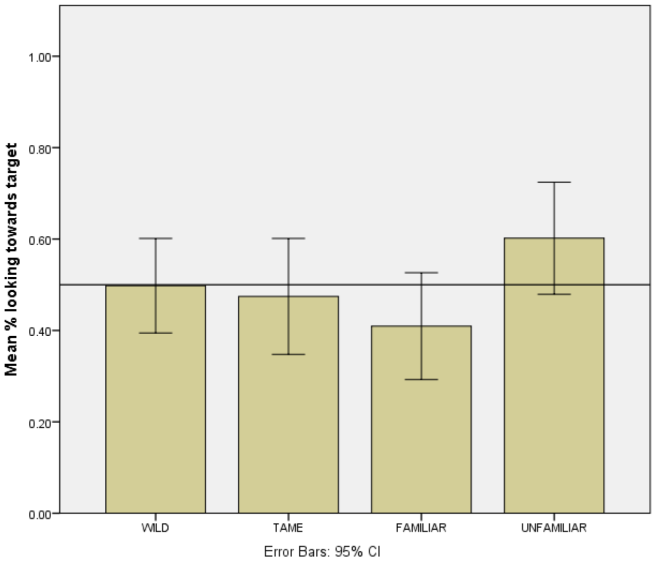

Figure 2: Results for CW0 infants

---

[2] Infant was scored as 0 if they had not yet produced the word form, whether or not they were reported to understand the form by the parent. Initially three scores were given (‘produces’, ‘understands’, ‘doesn’t understand’) but as no difference was found between ‘understands’ and ‘doesn’t understand’, these categories were merged.

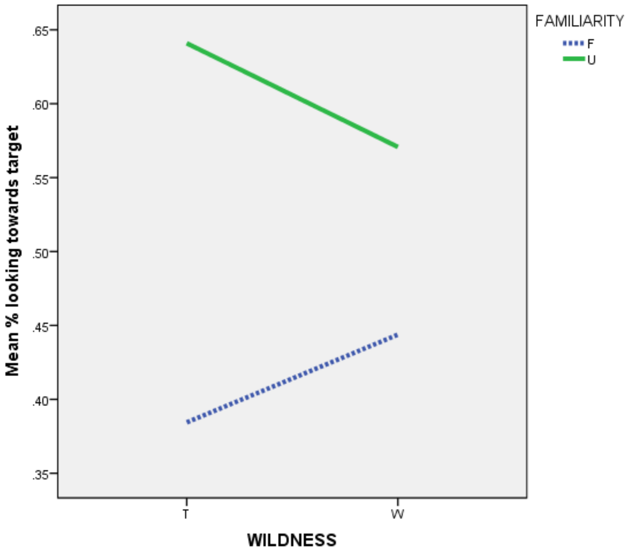

Figure 3: Effects of wildness and familiarity for CW0 infants

The OW0 group (n = 17) was then subjected to the same analysis, and a significant effect was found for familiarity (F(1, 16) = 5.346, p = .034), while wildness yielded a marginally significant effect (F(1, 16) = 4.125, p = .059). No effect was found for the interaction of wildness x familiarity. Again, Figures 4 and 5 show a bias towards L_U words, while infants in this group tended to respond above chance to T but not W stimuli.

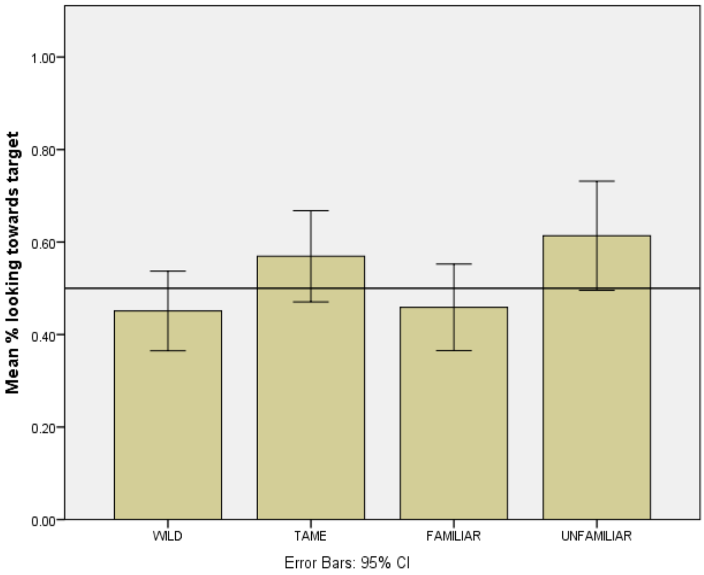

Figure 4: Results for OW0 infants

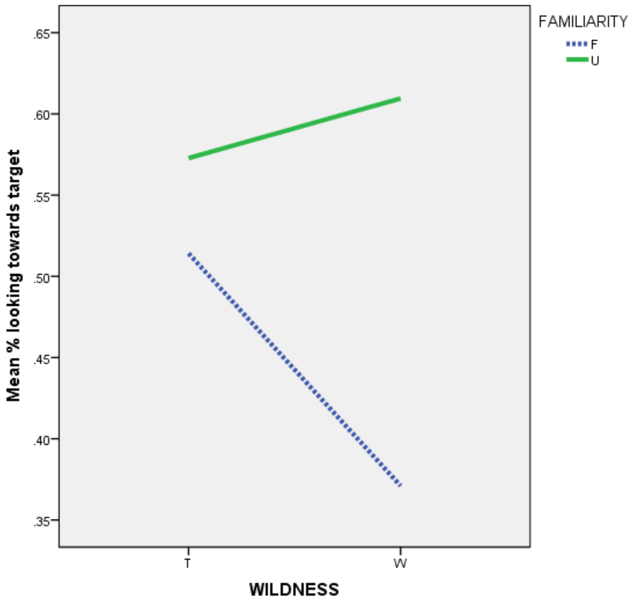

Figure 5: Effects of wildness and familiarity for OW0 infants

Finally, infants in knowledge group 0 for both CWs and OWs were analysed (n = 13), and the interaction between wildness and familiarity was found to be significant (F(1, 12) = 6.037, p = .03). The bias towards L_U stimuli can be observed in Figure 6, which was most pronounced in the W condition (Figure 7).

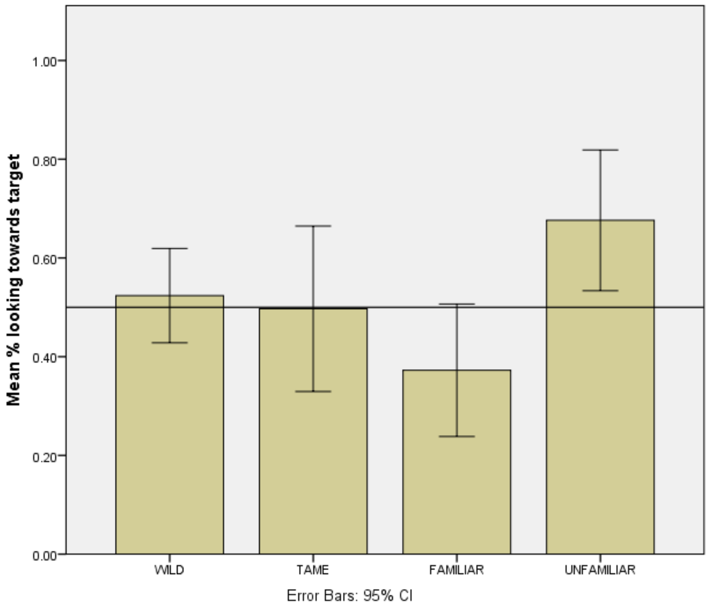

Figure 6: Results for OW0+CW0 infants

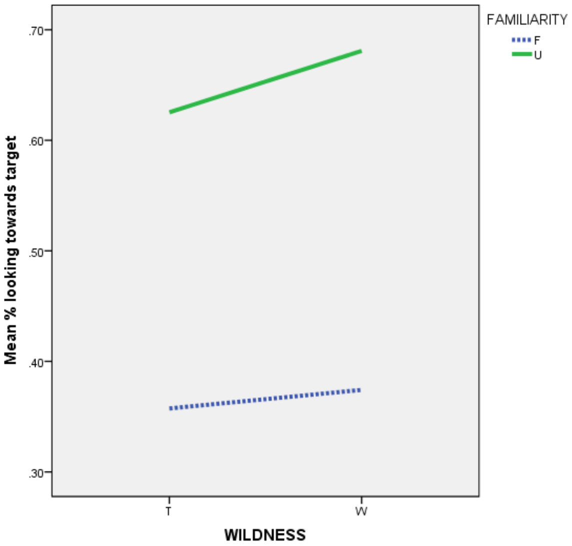

Figure 7: Effects of wildness and familiarity for OW0+CW0 infants

Results from this group show a consistent trend towards the L_U condition in all three parts of the analysis, demonstrating a bias towards this condition amongst infants in the earlier stages of language development. No trend relating to wildness was found in this condition, however, which indicates that wildness did not play a role in these infants' perception of the stimuli. This goes against the original hypothesis that infants will draw from the wild features of OWs in early language development in order to facilitate understanding.

3.2.2. Knowledge 1

Results for CW1 and OW1 groups were analysed using the same model. No significant effect was found for either group (CW1: n = 9, familiarity: F(1, 8) = .789, p = .4, wildness: F(1, 8) = .244, p = .635; OW1: n = 13, familiarity: F(1, 12) = .537, p = .478, wildness: F(1, 12) = .034, p = .857). In contrast to the results from knowledge group 0, it can be seen in Figure 8 that infants in the CW1 group tended to look to target above chance in all conditions except L_U. A similar effect can be found for infants in the OW1 group, who looked to target above chance in all conditions (Figure 9).

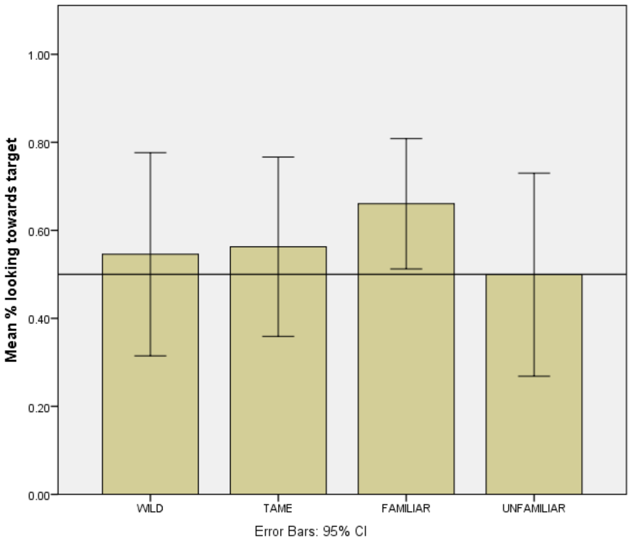

Figure 8: Results for CW1 infants

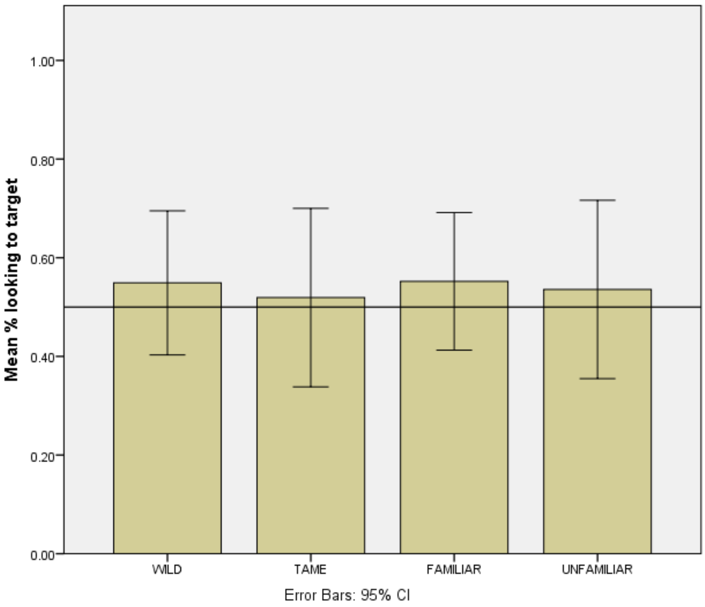

Figure 9: Results for OW1 infants

Infants with a knowledge score of 1 for both OWs and CWs were analysed (n= 7). No significant effect was found (Wildness: F(1, 6) = 1.764, p = .232, Familiarity: F(1, 6) = 3.995, p = .093), but Figure 10 shows biases towards both W and L_F stimuli. The lack of effect in these three tests could be due to low sample size, since the results show much longer looking times than those for knowledge group 0.

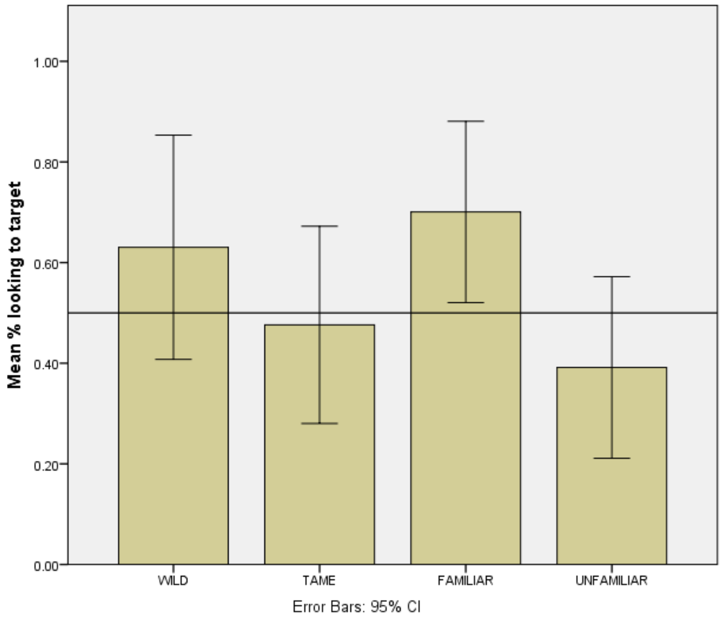

Figure 10: Results for OW1+CW1 infants

Infants in this group have been found to show a very different response to those in knowledge group 0. Figure 11 shows looking times for both OW+CW stimuli, where the difference between the groups can be seen more clearly. Results for responses to familiarity are almost inverted across the two groups, while wildness appears to affect responses only for infants in knowledge group 1.

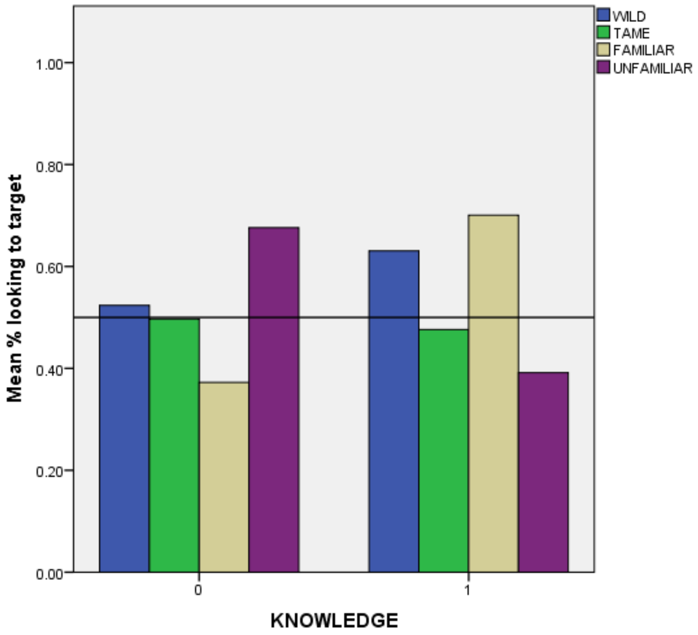

Figure 11: Results for both OW+CW knowledge groups across all conditions

3.3 *Reduplication*

Finally, the L_F and L_U stimuli were compared in order to discern whether any differences between forms in individual languages could be causing the discrepancy in response between the two groups. It was found that 100% of the L_U stimuli contained reduplication, of which all but one form were fully reduplicated (e.g. Chinese DOG [wʌŋwʌŋ], Arabic ROOSTER [kukuku:ku.kukuku:ku]), while only two of the six L_F OWs contain reduplication. These two stimuli (both L_F and L_U forms) were analysed separately in order to single out reduplication as a feature in this analysis. No preference was observed when the data were considered as a whole, while trends across the two knowledge groups remained consistent with previous findings (Figures 12 and 13): a significant effect was found for familiarity in knowledge group 0 (F(1, 10) = 5.248, p = .045 (n = 11)), and no effect was identified amongst the knowledge group 1 participants, possibly due to a low sample size for this group (n = 5).

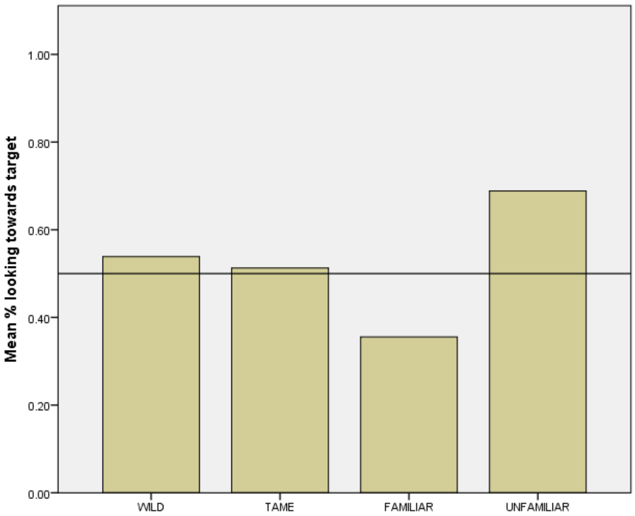

Figure 12: Results for reduplicated stimuli across OW0+CW0 infants

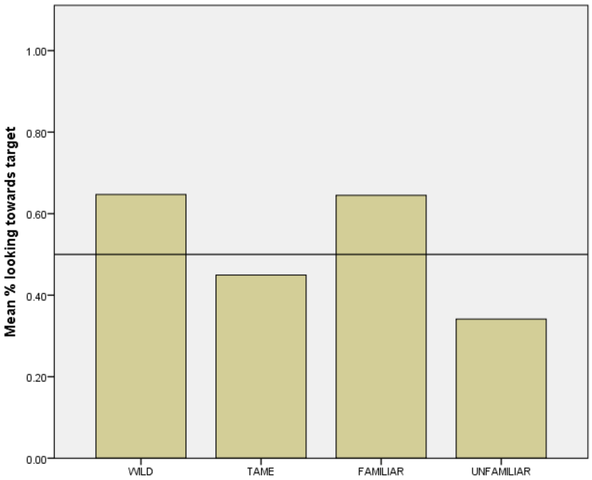

Figure 13: Results for reduplicated stimuli across OW1+CW1 infants

Duration of the individual stimuli was then measured using Praat, and a four-way repeated measures ANOVA showed L_U forms to be significantly longer then L_F forms (F(3, 8) = 14.61, p = .001). Furthermore, a two-way independent measures ANOVA found that W forms were significantly longer than T forms (F(1, 22) = 4.803, p = .039), as shown in Figure 14.

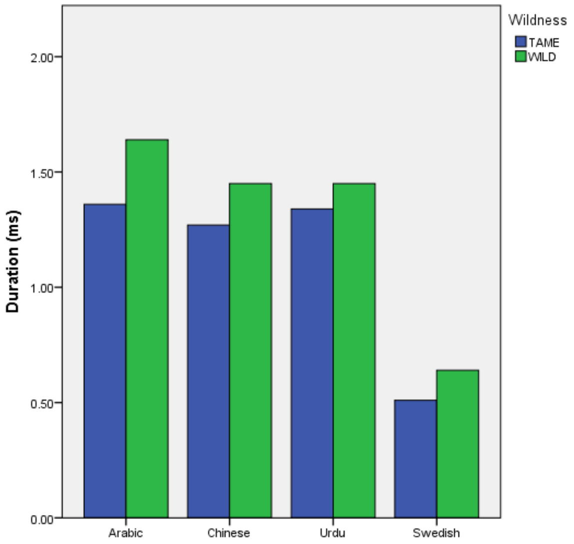

Figure 14: Average duration of W and T stimuli across languages

Results were then divided up according to duration of the various stimuli. Those stimuli that were shorter than 1.25s (the midpoint of the range of all results) were classed as 'short', and included all of the L_F and 11 of the L_U stimuli, and those which were 1.25s and longer were classed as 'long', and included only L_U stimuli. A two-way repeated-measures ANOVA showed that infants looked significantly longer at 'long' stimuli: F(1, 18) = 9.33, p = .007 (n = 19) (see Figure 15).

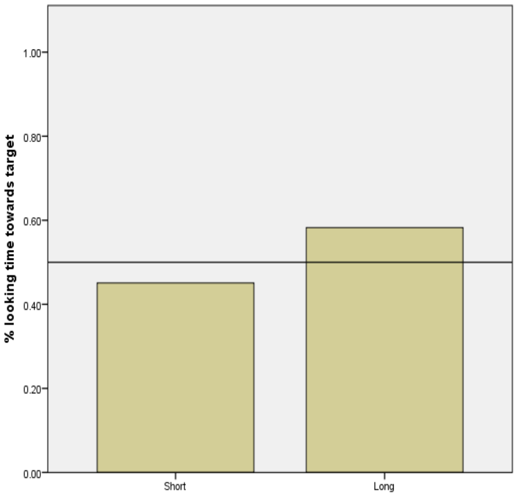

Figure 15: Average looking time towards 'short' and 'long' stimuli

While reduplication was not found to be a factor in its own right, it appears to bear some importance in terms of these results, which show that longer words elicit longer looking times. In the case of most of the individual stimuli used in this experiment, reduplication essentially doubled the infants' input, repeating the OWs in full to not only increase the amount of input received in each trial, but also to reiterate the information that the infants were receiving for each.

In sum, the results did not stand up to the hypothesis, and a closer analysis has shown that wildness does not appear to facilitate perception amongst infants who are unable to produce the word forms. Infants have been found to respond above chance to both W and T stimuli, but the results have shown the strongest biases towards L_U and L_F stimuli. It seems that wildness is a superfluous feature of OWs, which does not facilitate comprehension. Furthermore, reduplication has been found to be a prominent feature of the L_U stimuli, which may be an indirect contributing factor in the recognition of these word forms. Overall, the findings from this experiment suggest that infants who were able to produce the words perceived in the experiment attended to the speech stream differently to those who cannot yet produce the words, but that the infants as a group were biased towards those words which they perceived for the longest period. These responses will be discussed further below.

4. Discussion

These findings go against the hypothesis that infants would show the strongest response to W forms in both L_F and L_U conditions. No result was found when the data was considered as a whole, but patterns began to emerge in the results when these were broken down further to consider the infants’ knowledge of each of the target words. Infants who were not yet able to produce the target word form– either OW, CW or both–showed a significant preference for the L_U condition, while at no point did wildness appear to play a role for this group of participants. The opposite result was found for participants in knowledge group 1, who showed a preference for L_F stimuli throughout the data, but which was not substantiated by a significant result in the analysis.

The opposite effects observed in the results of knowledge groups 0 and 1 can explain why no effect can be seen in Figure 1: together, the two sets of results work to ‘cancel out’ one another, which reflects the extent of the discrepancy between the two groups. The obvious question to ask here is why infants at different stages of lexical development are producing these opposite effects, and what it might be that changes between the stages at which an infant comprehends a word form and produces a word form that leads them to this stark difference in perception.

Infants in knowledge group 0 responded most strongly to the L_U forms, indicating a strong reliance on the one thing that most of these forms had in common: extended word duration. The results seen here could reflect a lag in response time amongst this group, who may not have begun perceiving the word form as quickly as those who were able to produce the target word. Reduplication may also have contributed to these results, not only in adding duration to the infants’ experience of individual stimuli, but also by reiterating the phonological information for each of the L_U stimuli, and thereby facilitating perception. The use of reduplication in the stimuli used in this experiment is consistent with its common occurrence in IDS (Sundberg, 1998), which is thought to increase the salience of word forms in the early input.

While infants in knowledge group 1 do not show any significant trends in their results, emerging patterns in response to the L_F stimuli (both W and T) may relate to the influence of production on perception: infants in the early stages of word production are more likely to be drawn to those words in the input that they can produce themselves, and thus responses may be stronger towards these stimuli. This ‘articulatory filter’, as discussed by Vihman (1993), reflects the role of auditory feedback from an infant’s own output in terms of his perception of the input. As an infant’s phonological output develops, the articulatory filter prompts him to attend to those features of the input which are active in the output, a sort of ‘phonetic matching’ which supports the infant in the development of phonological memory (Vihman & DePaolis, 2000; Vihman, 2014). As stated by Vihman (2014), ‘the child’s first word production should *facilitate recognition of* (*and support attention to*)…words that resemble the word forms that are in the child's productive repertoire' (ch.9, emphasis added).This explanation can account for the increased attention to L_F forms in knowledge group 1's results, as phonetic representations are bolstered by the infants' production of these word forms. Again we see a familiarity response here, but this time this has been 'reset' to the least complex form–that is, the form that best fits the infants' output. Finally, the lack of consistent response to W or T stimuli suggests that these features may be arbitrary in relation to the phonological structure of the word form. It could also be that phonological wildness is specific to particular target words or even individual infants; more results are needed for this group if these options are to be investigated.

5. Conclusion

These results provide no concrete evidence for the role of wildness in word learning; wild features in the input appear to be arbitrary when paired with words the infant recognises and, more importantly, words he can already produce. However, the results do suggest that the reduplicative features of many OWs may prompt perception and, later, production in language development. It is likely that the wild vs. tame paradigm is specific to individual infants' experiences of individual word forms, and thus perception of OWs cannot be measured across-the-board in such a way. These findings highlight the essential interplay between perception and production in early language development, demonstrating the breadth of an infant's early lexical categories which account for both degeneracy and variability in the input, before the onset of production brings about perceptual narrowing, providing feedback specific to the phonological categories of the ambient language.

References

BEHRENS, H. 2007. "The input–output relationship in first language acquisition". *Language and Cognitive Processes*, 21(1-3), pp. 2-24.

BLOOM, P. 2000. *How Children Learn the Meanings of Words*. Cambridge,MA:MIT Press.

FENSON, L., DALE, P. S., REZNICK, J. S., BATES, E., THAL, D. J., & PETHICK, S. J. 1994. "Variability in early communicative development". *Monographs of the Society for Research in Child Development*, 59(5), Serial Number 242.

GENESEE, F., NICOLADIS, E. & PARADIS, J. 1995. "Language differentiation in early bilingual development". *Journal of Child Language*, 22, pp. 611-631.

IMAI, M., KITA, S., NAGUMO, M. & OKADA, H. 2008. "Sound symbolism facilitates early verb learning". *Cognition*, 109(1), pp. 54-65.

IVANOVA, G. 2006. "Sound-symbolic approach to Japanese mimetic words". *Toronto Working Papers in Linguistics*, 26, pp. 103-114.

IWASAKI, N., VINSON, D. P., & VIGLIOCCO, G. 2007. "What do English Speakers Know about *gera-gera* and *yota-yota*? A Cross-linguistic Investigation of Mimetic Words for Laughing and Walking". *Japanese-Language Education around the Globe*, 17, pp. 53-78.

KANTARTZIS, K., IMAI, M. & KITA, S. 2011. "Japanese Sound-Symbolism Facilitates Word Learning in English-Speaking Children". *Cognitive Science,* 35, pp. 575-586.

KAUSCHKE, C. & KLANN-DELIUS, G. 2007. "Characteristics of maternal input in relation to vocabulary development in children learning German". In I. Gülzow & N. Gagarina (eds.) *Frequency Effects in Language Acquisition: Defining the Limits of Frequency as an Explanatory Concept.* Berlin: Walter de Gruyter GmbH.

KITA, S. 1997. "Two-dimensional semantic analysis of Japanese mimetics". *Linguistics*, 35, pp. 379-415.

MENN, L. & VIHMAN, M. 2011. “Features in child phonology: inherent, emergent, or artefacts of analysis?”. In N. Clements & R. Ridouane (eds.), *Where Do Phonological Features Come From? The nature and sources of phonological primitives*. (Language Faculty and Beyond 6.) Amsterdam: John Benjamins.

RHODES, R. 1994. “Aural Images”. In J. Ohala, L. Hinton & J. Nichols (eds.) *Sound Symbolism.* Cambridge, UK: Cambridge University Press.

SUNDBERG, U. 1998. *Mother tongue – Phonetic Aspects of Infant-Directed Speech.* Stockholm: PERILUS.

TSUJIMURA, N. 2005. “Mimetic verbs and innovative verbs in the acquisition of Japanese”. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 31(1), pp. 371-382.

VIHMAN, M. M. 2014. *Phonological Development: The first two years*. (2nd ed.). Oxford: Blackwell.

VIHMAN, M. M. 1993. “Variable Paths to Early Word Production”. *Journal of Phonetics*, 21, pp. 61-82.

VIHMAN,, M. M. & DEPAOLIS, R. A. 2000, “The role of mimesis in infant language development: Evidence for phylogeny?”. In C. Knight, M. Studdert-Kennedy & J. Hurford (eds.) *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge, UK: Cambridge University Press.

Acknowledgements

I would like to thank Lena Renner, Ellen Marklund, Elisabet Cortes for their invaluable help in carrying out the eye-tracking study, as well as Professor Francisco Lacerda for allowing me to work at the University of Stockholm’s Babylab. This research could not have taken place without funding from the ESRC’s OIV scheme.

Catherine E. Laing
Department of Language and Linguistic Science
University of York
Heslington, York
YO10 5DD
United Kingdom
Email: catherine.laing@york.ac.uk