Acoustic Characteristics of Phonological Development in a Juvenile African Grey Parrot (Psittacus Erithacus) Who Is Learning Referential Speech

Acoustic Characteristics of Phonological Development in a Juvenile African Grey Parrot (*Psittacus erithacus*) Who Is Learning Referential Speech

Katia Zilber-Izhar

A Thesis in the Field of Biology

for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

November 2015

Abstract

Although young children can sometimes produce words in a near perfect form at a very early stage, several diary studies revealed that these correct first productions are usually followed by less faithful renditions, only to be returned later to relative accuracy. In order to investigate if this nonlinear pattern of children vocal production called "phonological regression" might also be shared with birds, we examined here the trajectory of vocal development of a young African Grey parrot (Athena) who is learning referential English. Parrots are excellent model systems for the study of speech acquisition as they possess advanced cognitive skills and are expert imitators of the human voice. By tracking Athena's acquisition of vowel-like sounds over the course of fifteen months using audio recordings and acoustic software programs, we analyzed her vocal development over time, from her first squeaks to her more distinct pronunciations, and compared her progress with human children and other parrots in the lab. Not one, but multiple U-shaped curves characterized her acquisition of isolated labels. Our results indicate that, like human children, parrots can experience the phenomenon of phonological regression.

Acknowledgments

I would like to first thank the members of my thesis committee, not only for their time but for their guidance and support. I am indebted to Dr. Irene Pepperberg who opened to me a window into the fascinating world of interspecies communication and kindly shared with me her wealth of experience in the field. Her contributions were paramount in my development as a scientist. I am also most appreciative of my thesis co-director, Dr. Bence Ölveczky, for agreeing to serve on this committee.

A very special thanks is due to Scott Smith, Athena's primary trainer. Without him, this work would not have been possible. His patience and special bond with Athena were fundamental in the success of the project.

I would also like to acknowledge Dr. James Morris from Harvard Extension School who assisted and guided me throughout the entirety of my thesis project.

Finally and most importantly, I would like to thank my husband, Lior. His support, patience, and encouragement were the foundations of my academic journey. I'd like to thank him for his faith in me and supporting me as I pursued my dreams.

Table of Contents

List of Tables

List of Figures

Chapter I

Introduction

Like humans, birds use vocalizations as a means to communicate and have evolved complex vocal systems. They use specific calls or songs to attract mates, repel rivals, claim territory ownership, sing a duet with a mate, beg for food, interact with the flock, reprimand an intruder or announce the presence of a predator (Marler & Slabbekoorn, 2004). For humans as well as many birds, communication is a learned behavior, and some common themes have emerged from the study of how birdsong and speech are acquired (reviewed in Bolhuis & Everaert, 2013; reviewed in Doupe & Kuhl, 1999; Marler, 1970b). Notably, the striking similarities between the ways that children learn to speak and birds learn to sing can provide direct insight into the developmental processes of human speech. One of these remarkable parallels is the existence of well-defined utterances that include a transitional period of private vocal practice (known as "subsong" in birds and "babbling" in humans). In songbirds, an initial memorization phase during which the song of a tutor is memorized in the form of a neural template is followed by a sensorimotor phase during which the young bird starts vocalizing and compares its own vocal output with the template (Konishi, 1965). These early, highly variable and crude vocalizations (subsong) are gradually modified and refined to match the adult template. Likewise, babies start with a period of close listening, and then transition to a babbling phase that precedes adult vocal production (reviewed in Bolhuis et al., 2010; Thorpe & Pilcher, 1958). Akin to subsong, babbling consists of long, rudimentary series of repeated

syllables, mostly voiced privately (Oller, 1986). The similarity between subsong and babbling had already been noticed by Darwin (1871) when he wrote in *The Descent of Man* that the first singing attempts [of songbirds] "may be compared to the endeavor in a child to babble" (mainly quoting the eighteenth-century vice president of the Royal Society, Daines Barrington).

Yet, during the noisy and unstructured subsong, some juvenile birds may occasionally produce mature versions of adult song patterns, well before they are supposed to display such singing capabilities. A few field reports support this intriguing observation and have documented species of wild passerine birds that produce sequences of songs matching the tutor model while still in the early phase of learning. For instance, when interacting with adults, juvenile white-crowned sparrows (*Zonotrichia leucophrys*) may sing fully developed, adult-like songs well before they reach sexual maturity (Baptista, 1983). Kroodsma (1974) described an 80-day old Bewick's wren (*Thryomanes bewickii*) that was exposed to song playback or a neighboring singing adult, and produced a developmental subsong characterized by adult-quality portions inserted between crude, ill-defined phrases.

The equivalent phenomenon, though more complex, has also been observed in human infants who are learning to speak. Individual children have been reported to produce an early word in a near perfect form while still in the babbling phase. Interestingly however, usually around the time they have acquired their first fifty words, their phonology "regresses", only to return, much later, to a more advanced form. In these cases, some isolated first words, whose renditions were surprisingly accurate for several months, show suddenly a loss of correct production, before being articulated correctly

again. Several longitudinal studies and diaries have documented this phenomenon. The classic case is the famous production of "pretty" by Leopold's daughter, Hildegard (Leopold, 1939, 1947). "Pretty" was Hildegard's first stable word that she pronounced with near perfect accuracy at the age of ten months. Then, however, at the age of eighteen months, it gave way to "pittee" [pɪti] and a month later to "biddee" [bɪdi]. Another often cited example is the extraordinary treatment of "turtle" by Nicholas, the son of Peter and Jill de Villiers (de Villiers & de Villiers, 1979). At fifteen months, Nicholas produced a perfect "turtle" which became at eighteen months "kurka". Although Nicholas was able to pronounce the components syllables of "turtle" correctly, he would no longer pronounce the whole word. Bleile & Tomblin (1991) reported another case of phonological regression whereby a two-year old boy named Jake would, over a short period of time, no longer produce a sound in a newly acquired word. After Jake had learned "thunder" and had articulated it in a perfect fashion with an initial "th", he changed it to "sunder" and lost his accurate pronunciation of the initial phoneme "th" [θ]. Johnson and Reimers (2010) witnessed the case of Amy, whose early pronunciation of "juice" was later reduced to "dus" when her vocabulary increased. Similarly, Alice, also the subject of a phonological development study, produced at ten months and for a period of five months early versions of "hi" whose accuracy exceeded subsequent renditions such as [ha:ji] (Vihman, 1992).

In most cases, regression is explained by the emergence of rules and systematicity for the pronunciation of words when children start acquiring many new words, very quickly (Anisfeld, 1984; Bleile & Tomblin, 1991; de Villiers & de Villiers, 1979; Leopold, 1947; Vihman & Kunnari, 2006). Indeed, once they have acquired a small

lexicon of early words, children pick up and generalize production patterns of the words they use most. At the same time, they extend these emergent output patterns to a wider range of adult word shapes. As vocabulary expands, words become more similar to one another (Vihman & Kunnari, 2006). For instance, a nineteen-month-old French girl named Beryl showed a strong <aCV> pattern that she over-imposed on words that came to resemble one another: [afɔ] for "éléfant" (elephant), [ato] for "bateau" (ship), [alo] for "agneau" (sheep) and [aço] for "cerceau" (circle) (Wauquier & Yamaguchi, 2013). However, the beginning of systematicity and organization is often accompanied by a decrease in production accuracy. Children project their developing vocal schemes onto adult word forms that require radical changes to fit the emergent phonological patterns ("word templates"), leading to less accuracy in terms of matching the adult models (Vihman & Kunnari, 2006). Leopold (1947) invoked such rules to explain the dramatic change and the regression in accuracy in Hildegard's "pretty" when the word lost the cluster "pr" to fall in line with more regular "rules of substitution". Similarly, Vihman and Kunnari (2006) described the emergence of word templates that are accompanied by the loss of correct pronunciation through the case of a fifteen-month-old French boy (Charles), who followed a pattern of first consonant omission to bring into line some target sequences. For instance, he produced [apo] for 'chapeau" (hat), [apa] for "lapin" (rabbit) and [apa] for "va pas" (doesn't fit). The pressure to assimilate non-matching adult productions to fit the emergent child's output patterns is responsible for the decrease in overall accuracy but also the increase in the rate of lexical learning and inner coherence among the child's own forms (de Villiers & de Villiers, 1979; Vihman, 2014). At the same time, the phonetic development that is taking place in the child's

phonological system permits an eventual return to accuracy in reproduction of adult forms.

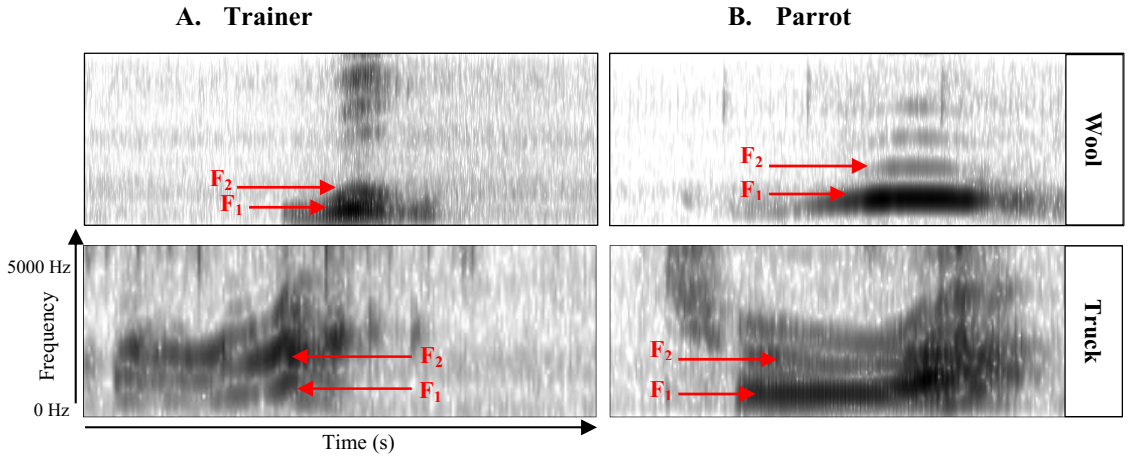
Although the studies in both songbirds and humans may indicate potential similarities between the two groups with regard to the phenomenon of phonological regression, the comparisons cannot lead to any clear conclusion concerning vocalization acquisition patterns in general because songbirds lack some of the crucial characteristics relevant to the speech faculty (e.g. referentiality). Parrots, in contrast, seem to be better suited models. Like humans, they are open-ended learners, with the ability to acquire new vocalizations throughout their life. As do children in the early stages of speech learning, they also engage in "sound play" that includes playing with the combinations of phonemes to create new sound patterns (Pepperberg et al., 1991). They produce highly complex calls and are adept at imitating heterospecific sounds, whether it is other birds in the wild (Cruickshank, Gautier, & Chappuis, 1993) or human words in captivity (Pepperberg, 1999). Whereas some other birds can also mimic words, such as mynahs (Klatt & Stefanski, 1974) and corvid songbirds (Petkov & Jarvis, 2012), parrots can actually make complex use of human speech. In laboratory studies, African Grey parrots have demonstrated human-like ability in many aspects of their use of speech. This includes: understanding the connection between words and what they stand for in real life, an ability considered key to aspects of language learning (referential communication, Pepperberg, 2006), recombining individual phonemes in novel ways to create new labels (vocal segmentation, Pepperberg, 2007), and using words with varying social contexts (Colbert-White, Covington, & Fragaszy, 2011). Grey parrots are also renowned for their intelligence and advanced cognitive aptitudes (Giret et al., 2011; Pepperberg, 1999).

Alex, the famous parrot, could identify correctly the number of a subset of items of a given category and color presented among an heterogeneous collection (Pepperberg, 1994). Burish et al. (2004) reported that African Greys belong to the five species with the largest telencephalon ratio, out of a list of 154 bird species. These complex cognitive skills challenge the aptitudes of the great apes in many domains (Emery, 2006; Pepperberg, 1999; Pepperberg & Carey, 2012). Parrots also exhibit rich and complex social behavior. In the wild, they use their calls to sing duets and form long-lasting pairs, and at home or in the laboratory, they often establish strong social bonds with their human caretakers (Colbert-White, Covington, & Fragaszy, 2011; May, 2004). Taken together, these findings make the African Grey parrot an exceptionally interesting candidate for the study of speech acquisition in general. Whereas many aspects of functional use of speech have been documented in African Grey parrots (Pepperberg, 1999), the phenomenon of phonological regression and the development of speech patterns in a young parrot have never been studied before.

The aim of the present study was to investigate the vocal development of a young female African Grey parrot ("Athena") in a laboratory setting. In particular, this research investigated the acoustic pattern of phonological development of a parrot who is learning to communicate referentially in English. We attempted to determine if Athena experienced a nonlinear advance (in the form of a U-shaped curve) in the acquisition of selected English labels, as children occasionally do. We hypothesized that, if a human infant, who is learning to speak, and an African Grey parrot, who is learning to communicate referentially in English, follow a similar phonological U-shaped trajectory, then these data would imply that this similar pattern of speech development called

"phonological regression" has an evolutionary basis and would have evolved at least twice, once in parrots and once in humans. To evaluate this idea, we monitored the variations of the formant frequencies of vowel sounds contained in English labels recorded from Athena over a course of fifteen months. "Formants", which reflect the resonances of the vocal tract, are concentration of acoustic energy around a particular frequency in the speech wave and are displayed as dark bands on spectrograms (Figure 1).

Figure 1: Spectral comparisons of the speech utterances "wool" and "truck"



This figure represents two wideband spectrograms illustrating the trainer (A) and a parrot (B) saying the labels "wool" and "truck". The first two formants, F_1 (formant 1) and F_2 (formant 2) are indicated. In the spectrogram of "wool", the other formants, F_3 and F_4, though not shown, are clearly visible.

Formants are particularly important as their patterns provide the acoustic cues essential for the characterization of vowels. For instance, the difference between the vowel tokens in the labels "heed" and "hat" (corresponding to the vowels /i/ and / æ / of the International Phonetic Alphabet (IPA)) is based on different formant frequencies only.

African Grey parrots use the two-chamber structure of their vocal tract to allow selective resonance of the sound generated by the syrinx in combination with changes in the trachea length and oropharyngeal cavity, as well as unique lingual articulations, to produce vocalizations with vowel-like and consonant-like qualities (Warren, Patterson, & Pepperberg, 1996). Such a speech-like formant system contributes to their talent as imitators of the human voice (Beckers, Nelson, & Suthers, 2004; Bottoni, Masin, & Lenti-Boero, 2009). Although parrots' absolute formant frequencies may differ from human values, the relative changes in formant frequencies during language acquisition follow the same trend and therefore can be compared. It is worth noting that Athena had acquired no words prior to starting this experiment and that the recordings covered her attempts at learning her first labels. We examined the trajectory patterns of Athena's vowel frequency curve in reference to her trainer's fixed formant values. Athena's formant frequencies were also tracked against her trained lab parrot companion (Griffin), as he had been used in modeling sessions and may have had an influence on her speech patterns. Finally, we traced the development of the label duration and looked for potential correlations with the formant patterns.

Chapter 2

Materials and Methods

Subject

Subject was a juvenile female African Grey parrot, Athena, five months at the beginning of the experiment. She had been hand-raised and was purchased from a pet store one month previously. During the day, the bird lived in a laboratory setting, atop of her cage or on parrot stands; during the night, she was housed in an aviary in the animal care facility in a standard cage (90x50x80 cm). Water, parrot pellets and dried pasta were available *ad libitum*. The parrot was also fed fresh fruits, vegetables and grains three times a day. It is worth noting that another Grey parrot – a nineteen-year-old male ("Griffin") – was also present in the lab at all times. Although Griffin was not involved in the present experiment, except occasionally for modeling purposes, he had been the subject of continuing studies on interspecies communication and had already labels for many objects, including several types of toys and materials being used in the research with Athena. In contrast, Athena had received no formal training prior to these experiments and had acquired no human vocalizations.

General Procedure and Training

To determine the pattern of variability of Athena's vocalizations while under development, we used English vowel-like sounds recorded from her when she attempted to pronounce specific labels, measured their formant frequencies and traced their

patterns. She was recorded three to four times per week between October 2013 through December 2014, during elicited recording sessions (95% of recorded vocalizations) and while producing spontaneous calls or "babbling" (5% of recorded vocalizations). Two additional recording sessions took place in March 2015. During trained recording sessions, while the bird watched and listened, the primary trainer held an object and asked questions about them (e.g., "What toy?", "What matter?"), encouraging Athena to vocalize. If she did not respond, he said the label. The word was repeated several times with a slightly high pitched voice and an exaggerated intonation, marking clear pauses between repetitions. Recurrent sentence frames such as "That's a ….!" or "You're chewing a …", where the target word that enters the frames is usually heavily stressed, were used to draw Athena's attention to this label. In addition, the label was usually placed at the end of the sentence, because parrots, like humans, tend to pick-up information at the end of phrases (Pepperberg, 1999). Spontaneous recordings were made when Athena was left alone by her trainer for several minutes during trained recording sessions.

In one-on-one recording sessions with her primary trainer, new words were slowly introduced to Athena. However, to increase the pace of her lexical acquisition, Athena started in January 2014 a tutoring protocol called the model/rival (M/R) procedure that had been developed by Todt (1975) and was further adapted by Pepperberg (1981). This new training approach took place during or in addition to the recording sessions. In brief, M/R training involves a three-way interaction between two human tutors and the avian pupil. The purpose of the training was to introduce new labels and concepts, but also to help in correcting pronunciation. Typical sessions begin with

the bird watching two humans holding an object. One of them acts as a trainer while the other acts both as a model for the bird's response and as a rival for the object and trainer's attention. The trainer questions the model/rival about the item ("What toy?", "What matter?", "What do you want?"). Praise and the object itself are given as the only rewards for the correct answer, thus reinforcing the association between the referent and the label to be learned. The model/rival also occasionally produces errors (incorrect responses or mispronunciation) which are punished by showing disapproval and scolding; the object is also removed from view. The interaction is repeated by reversing the roles of the human trainers, so that the parrot sees that one person is not always the questioner and the other the respondent, and a correction procedure takes place. The parrot is then engaged in the exchange, being questioned and rewarded for attempts at a correct response, or reprimanded for errors. The M/R technique was used several times a week. Griffin, the other parrot in the lab, sometimes was used as a model for Athena. The length of each session depends on the attention's span of the bird, which in the case of Athena, rarely exceeded ten-fifteen minutes. Athena was trained on several tokens that belonged to two categories: toys (eight labels) and matter (eight labels).

Acoustic Data Collection and Analysis

Database Construction

The recording sessions covered four periods. The first four months (October–January 2014) were dedicated to collect baseline data that was exclusively composed of calls, whistles and amorphous sounds. With the introduction of the M/R training protocol and extensive one-on-one sessions with her trainer, Athena started producing vowel-like

sounds that were recorded over the next eleven months. Recordings were made using a Sennheiser microphone directly into a MacBook laptop and later into a HP Envy laptop. Unfortunately, the recordings made with the HP Envy laptop turned out to be of poor quality and, therefore, we switched back to the MacBook. This technical recording problem resulted in a loss of data over the period of June 17, 2014 through September 30, 2014. Although many recordings were unusable, we were nevertheless able to restore and analyze a small number of vocalizations. The last two recording sessions in March 2015 represented the endpoints. Recordings were digitized at 16 bits at a 44100 Hz sampling frequency and saved as AUP files. Out of Athena's lexicon-in-progress, we selected the five tokens in which she was most interested, hoping this decision would speed-up the learning process: "wool", "wood", "paper", "nylon" and "truck". For each session, the best approximations of each label were then extracted and converted into a WAV format using Audacity (version 2.0.6, retrieved from http://audacity.sourceforge.net/) for further spectrographic analyses. 129 sessions were recorded for a total of 7134 minutes. We obtained 1131 parrot vocalizations across the five selected labels, which were used to analyze the following four English vowels: 514 [ʊ] like in "w<u>oo</u>l" or in "w<u>oo</u>d"; 199 [ə] like in "pap<u>e</u>r"; 228 [ɑ] like in "nyl<u>o</u>n" and 190 [ʌ] like in "tr<u>u</u>ck" (Table 1). Similarly, the five selected labels were recorded from Athena's primary trainer in a flat, calm voice: 27 "wool" and 31 "wood", totaling 58 samples of human [ʊ] vowel, 23 "paper" and 23 "truck" for samples of human [ə] and [ʌ] vowels, and 20 "nylon" as samples of human [ɑ].

Table 1. Labels selected for spectrographic analysis

Label	Category	Object	IPA Vowel	Number of Samples
Wool	Matter		[ʊ]	195
Wood	Matter		[ʊ]	319
Paper	Matter		[ə]	199
Nylon	Matter		[ɑ]	228
Truck	Toy		[ʌ]	190

This table lists and describes the five labels selected for the spectrographic analysis.

In addition, because the bird was mostly exposed to the trainer's exaggerated intonation and high-pitched voice, a second set of his vowel sounds was used as a control and thirty samples of each label said during the training sessions were extracted randomly for further acoustic comparison with the same labels said alone, in a normal tone of voice. To perform the comparative analysis between Athena and her parrot companion, we recorded Griffin naming the targeted labels, in separate sessions.

Acoustic Analysis

The various speech sounds (of Athena, the trainer and Griffin) were analyzed using the Praat software version 5.4.04 (Boersma, 2001). Acoustic analysis involved obtaining the first two formants (F_1, F_2) of the studied vowels of the parrot and human model. The analysis focused on the second formant (F_2) because the first formant (F_1), which is an indication of tracheal change, varies little across vowels (Patterson & Pepperberg, 1994) In contrast, F_2, which is produced by tongue articulations, beak opening, glottis and larynx changes, varies significantly and is a good correlate of vowel identity (Warren, Patterson, & Pepperberg, 1996). The formant frequencies of each vowel were obtained by selecting with interactive cursors the appropriate portion of the vowel and then querying the program for the mean value of each formant over that range. In the present analysis, for both human and parrot, the formant tracking system was instructed to identify five formants over the range of 0–5,500 Hz using a frequency window of 25 ms and a dynamic range of 20 dB. Parrot non-speech vocalizations that consisted of chirps, whistles or squawks were included as long as Praat's formant tracking system was returning values. The length of all sounds was also measured and reported. To visualize

the progress in Athena's vocal vowel expression, a calculated relative ΔF_2 score was used. The relative ΔF_2 score represents the relative difference in F_2 values between Athena and the trainer in each given time point normalized to the trainer's F_2 value. This calculated score is expressed by the formula

$$r\Delta F2 = 1 - \frac{|F2A - F2T|}{F2T}$$

(where F2A=raw F_2 value of Athena, F2T=raw F_2 value of the trainer)

A similar approach was used to visualize Athena's utterance duration progress with respect to her trainer.

Chapter 3

Results

Vowel-like Sound Development in an African Grey Parrot

Control against Trainer's Normal Tone of Voice Used as a Reference for Acoustic Analysis

In order to assess the developmental pattern of Athena's speech, we compared her F_2 formants produced in the different time points of the recording period with the F_2 formants values produced by her trainer. T-tests on the vowels [ʊ], [ə] and [ɑ] did not reveal a significant difference between the mean F_2 formant values of the trainer when he spoke in a normal, flat voice and when he spoke with an exaggerated intonation during the recording sessions (Table 2). Therefore, for consistency, we performed the comparative analysis for all vowels using the mean F_2 values of the trainer that were obtained from separate recordings (Table 3).

Table 2. Means of F_2 and number of samples for trainer in separate recordings ("alone") and trainer in recordings with Athena, and t-test on F_2 of trainer in both conditions

	[ʊ] as in w<u>oo</u>l		**[ə] as in pap<u>e</u>r**		**[ɑ] as in nyl<u>o</u>n**	
	F_2	**Nb. of samples**	**F_2**	**Nb. of samples**	**F_2**	**Nb. of samples**
Trainer "alone"	815 (± 137)	29	1283 (± 129)	23	1070 (± 449)	20
Trainer with Athena	830 (± 170)	30	1262 (± 113)	30	1061 (± 108)	30
t-test on F_2	p=0.72		p=0.54		p=0.93	

This table includes mean formant F_2 values of Athena's trainer in two conditions: during the recordings sessions with her and in separate recordings. During recorded training sessions, an exaggerated voice was usually used whereas while alone, the labels were spoken in a flat tone. For each vowel and each type of recording, the number of samples is indicated. For each vowel, a t-test was performed on the F_2 frequency value of the trainer recorded in both conditions. Standard deviations are listed in parentheses.

Table 3. Means of F_2 of trainer recorded in separate conditions

	[ʊ] as in w<u>oo</u>d and w<u>oo</u>l	**[ə] as in pap<u>e</u>r**	**[ɑ] as in nyl<u>o</u>n**	**[ʌ] as in tr<u>u</u>ck**
	F_2	**F_2**	**F_2**	**F_2**
Trainer	980[a] (± 217)	1283 (± 129)	1070 (± 449)	1129 (± 158)

[a] F_2 is the mean of 815 (wool) and 1145 (wood).

This table lists the mean formants values of Athena's trainer that were used as reference in the vocal development analysis. Standard deviations are listed in parentheses.

Comparison of Athena's Mean Formant Values with the Corresponding Human Formants

Mean formant values for Athena, her trainer and human males across vowels are provided in Table 4.

Table 4. Means of F_2 for humans and an African Grey parrot across vowels

	[ʊ] as in w<u>oo</u>d, w<u>oo</u>l	**[ə] as in pap<u>er</u>**	**[ɑ] as in nyl<u>o</u>n**	**[ʌ] as in tr<u>u</u>ck**
	F_2	**F_2**	**F_2**	**F_2**
Athena[a]	1124[b] (± 145)	1194 (± 63)	1146 (± 72)	1393 (± 204)
Trainer	980[c] (± 217)	1283 (± 129)	1070 (± 449)	1129 (± 158)
Human males[d]	1020	1400[e]	1090	1190

[a] Athena's mean values were computed from the endpoint recordings, when the labels were recognizable.

[b] F_2 is the mean of 1135 (wool) and 1113 (wood).

[c] F_2 is the mean of 815 (wool) and 1145 (wood).

[d] Values are from Peterson and Barney (1952).

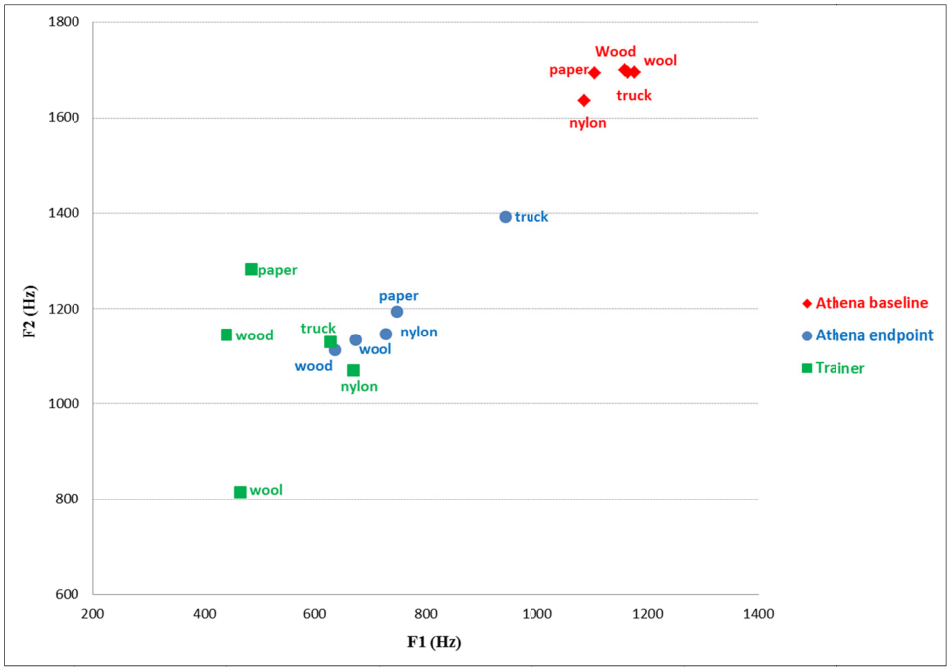
[e] Value is from Lindblom (1986) because it was not available in Peterson and Barney (1952).

This table shows the mean values of formant F_2 and standard deviations for Athena, Athena's trainer and human adult males across the four vowels that were contained in the selected labels. This table also shows that the trainer's values are comparable with data published in the literature. Standard deviations are listed in parentheses when available.

Athena's mean F_2's range across vowels from 1124–1393 Hz compared to 980–1283 Hz for her trainer and 1020–1400 Hz for the reported sample of men. Athena's mean values are restricted in regard with her trainer: her range of frequency values for F_2 covers only 52.5% of her trainer's. In terms of absolute formant values across vowels, there are similarities and differences between avian and human values. Athena's second formant differs from her trainer's corresponding formant by less than 10% for [ə] (6.9%) and [ɑ] (7.1%). The resemblance is weaker with [ʊ], though still considerable, as her data differ from her trainer's by less than 15% (14.7%). The difference is most striking with respect to Athena's F_2 for [ʌ], as it differs from that of her trainer by almost 25% (23.4%).

The previous analysis gives us only a "snapshot" of the parrot's second formant frequency at the end of the experiment with respect to the corresponding formant of the trainer. In contrast, a map of the vowel distribution of Athena at baseline and endpoint provides evidence of a clear trend towards matching her trainer's formant space (Figure 2). The shift from her initial clustered position at baseline towards the formant space of the human model at endpoint appears clearly on the plot.

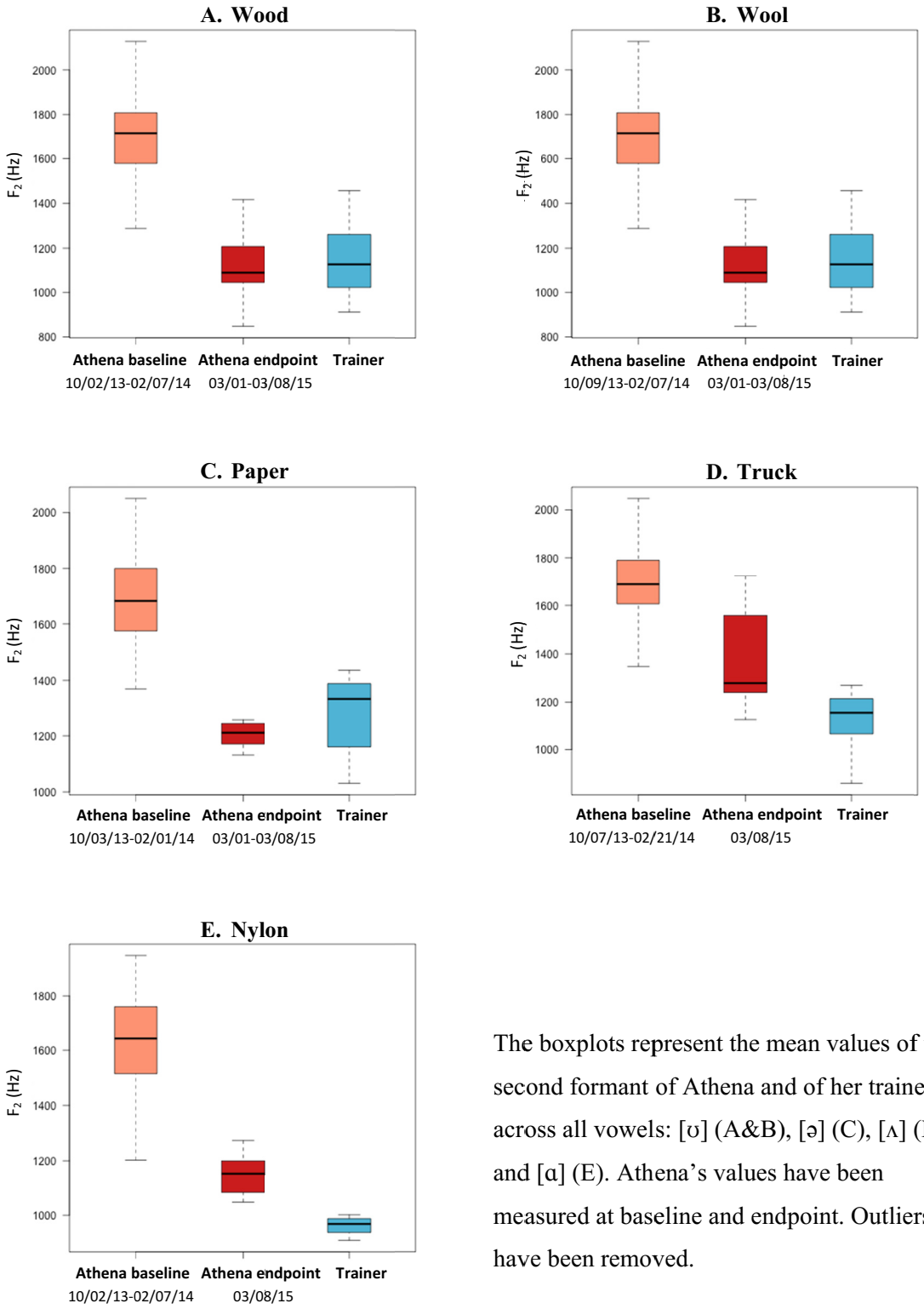
Figure 2. Vowel distribution map of Athena and her trainer



This graph plots the mean frequencies of F_1 and F_2 of the trainer and Athena for each selected label (i.e. the vowel contained in that label). In particular, this graph illustrates the formant space occupied by Athena at baseline and endpoint, and the transition towards the trainer's vowel space at the end of the experiment.

These findings are consistent with a more detailed comparison of Athena's mean F_2's at baseline and endpoint and her trainer's (Figure 3).

Figure 3. Boxplots of the mean F_2 across all vowels of Athena and her trainer



The boxplots represent the mean values of the second formant of Athena and of her trainer across all vowels: [ʊ] (A&B), [ə] (C), [ʌ] (D) and [ɑ] (E). Athena's values have been measured at baseline and endpoint. Outliers have been removed.

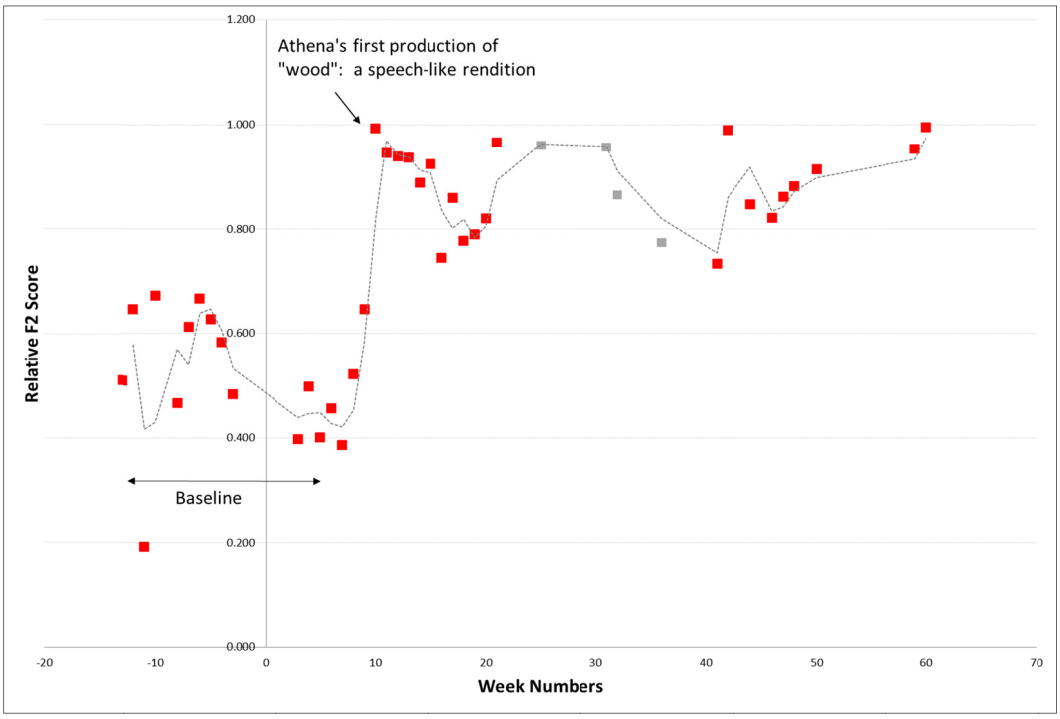
Athena's median values for all labels are significantly much lower at endpoint than at baseline (two-tailed t test; "wood", p= 6.78E-14; "wool", p= 1.99E-08; "paper", p= 1.68E-18; "truck", p= 0.001; "nylon", p= 1.24E-14) and the spreads are also narrower. Athena's precise imitation of the acoustic characteristics of her trainer when producing the vowel-like sound contained in the label "wood" is revealed by almost exact matching of the median lines in Figure 3A (Athena's mean F_2=1113; trainer's mean F_2=1145). Although not all the distributions overlapped (e.g. wool, nylon), the boxplots highlight the parrot's overall effort towards creating an accurate imitation of the formant frequencies of her human target. Interestingly, a one-way ANOVA test revealed that there was no significant difference between the mean frequencies at baseline across her five labels at the $p<0.05$ level ($F(4, 287)=1.93$, $p=0.106$). All the more so, the mean F_2 values of her "wool", "wood" and "paper" were equal or almost equal (1695 for "wood" and "wool", 1693 for "paper").

Athena's Vocal Development and F_2 Formant Matching with her Trainer - Acoustic Analysis per Label

In the following sections, we traced Athena's developmental path from baseline to endpoint over a period of fifteen months and analyzed the acoustic characteristics of each vowel-sound contained in the five tokens that we selected.

Production of "wood". Athena first produced "wood" on 3/4/2014 (week 10) in the presence of her trainer. This initial rendition was surprisingly clear and sounded human-like, with a relative F_2 score of 0.994. The developmental pattern for such an accurate and early production shows an exponential learning curve that began at week 7 (Figure 4).

Figure 4. Athena's mean F_2 of the vowel [ʊ] as in "w<u>oo</u>d" in reference to her trainer's mean F_2 (10/2/2013–3/8/2015)



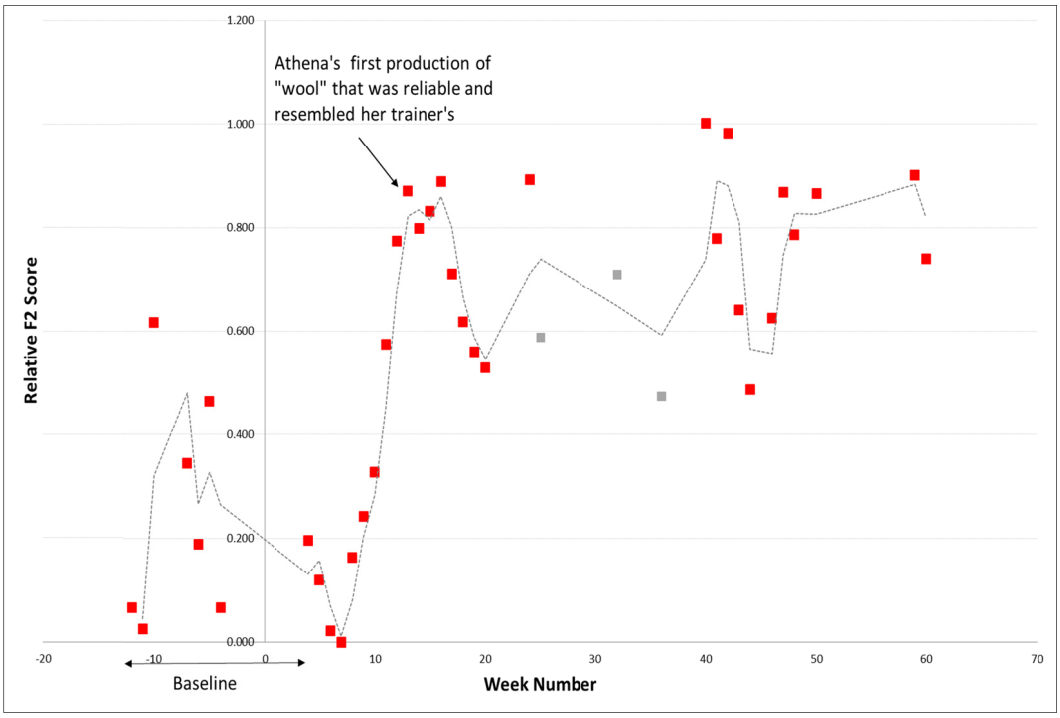
This graph represents the frequencies values of Athena's F_2 against her trainer's for the vowel [ʊ] as contained in "w<u>oo</u>d". Outliers have been removed according to the modified Thompson tau technique. Data points in grey represent the period of data loss/poor recordings.

Athena's level of pronunciation remained excellent for about six weeks, but then plummeted during the third week of April (week 16, relative F_2 score=0.745). It started improving again a month later, roughly around week 20, stabilized for a short period at an average score of 0.962 and then deteriorated again over the entire summer, reaching its lowest point since she had pronounced her first utterance at a relative F_2 score of 0.734. At that time, squeaks and squawks that constituted the vocalizations of the baseline became prominent again. However, the productions returned to be more accurate and reliable during the second week of October, reappearing in a form similar as the early accuracy experienced in March (relative F_2 score=0.990). This period was followed by a

short decline in production quality but Athena recovered immediately. The whistles became less frequent and her utterances more closely resembled a child's attempts to speak than a parrot's calls. From week 46 until the last measurement (performed on 3/8/2015), she had been constantly improving. She reached again a clear and reliable level of speech when we measured the two endpoints in March 2015, which scored 0.954 and 0.996.

Production of "wool". "Wool" followed similar patterns and timetable of development as "wood" (Figure 5).

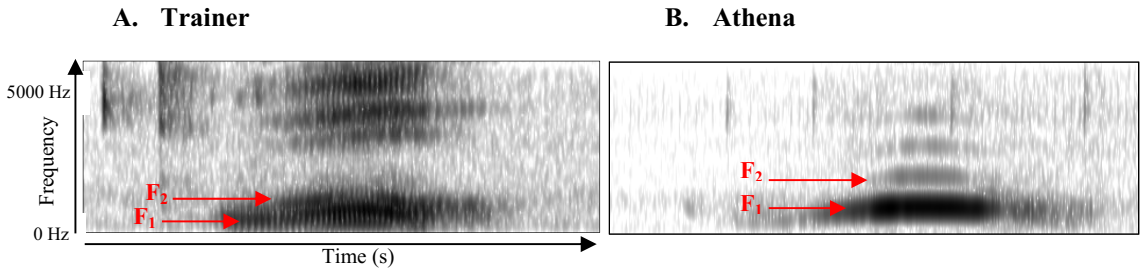
Figure 5. Athena's mean F_2 of the vowel [ʊ] as in "w<u>oo</u>l" in reference to her trainer's mean F_2 (10/9/2013–3/8/2015)



This graph represents the frequency values of Athena's F_2 against her trainer's for the vowel [ʊ] as contained in "wool". Outliers have been removed according to the modified Thompson tau technique. Data points in grey represent the period of data loss/poor recordings.

The results of the trend analysis confirm a sharp learning increase that began at week 7 and reached a peak during week 13. This change (from a relative F_2 score of 0.0003 at week 7 to 0.8708 at week 13) coincided with a considerable improvement in pronunciation of the label which became clear and reliable. The comparison of spectrograms computed at that time and representing the word "wool" by Athena and her trainer support the claim that she precisely imitated the utterance (Figure 6).

Figure 6: Spectral comparison of the speech utterance "wool" during week 13



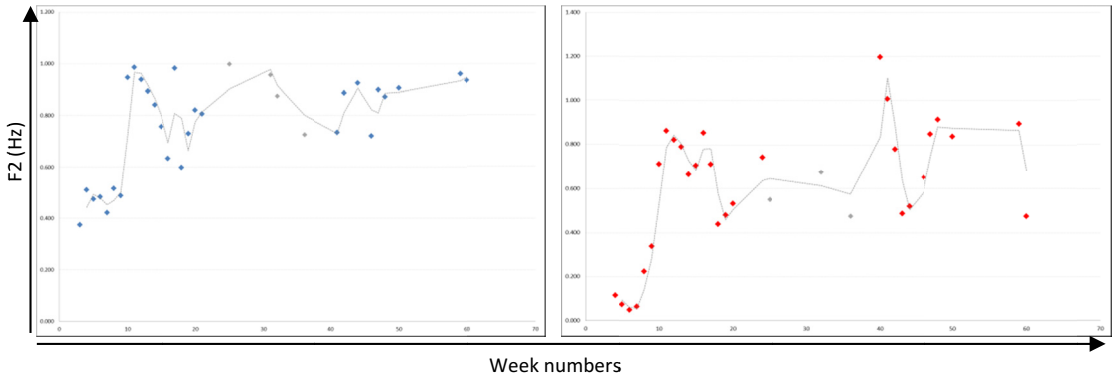
These two wideband spectrograms illustrate the label "wool" of the trainer (A) compared to Athena (B). The first two formants, F_1 and F_2, are indicated. Both recordings were made on the same day (3/15/2014), during week 13. Note the similar shapes between the two images.

This initial exponential growth was subsequently followed by a slow-down, and Athena's vocal productions for "wool" stabilized for a few weeks at an average relative F_2 of 0.847. However, as was reported with "wood", week 17 marked the onset of a loss of accuracy that persisted until week 20 (relative F_2 score=0.531). Likewise, a small improvement in June (weeks 24, 25) was overcome by an overall decrease over the summer, reaching an all-low relative F_2 score of 0.474 during week 36. However, again, the same trend observed in the vocal development of the label "wood" after week 40 occurred in Athena's production of "wool": a dramatic increase in her relative F_2 score led to a change in her pronunciation which suddenly resembled that of her trainer. Yet,

this near perfect human production of "wool" was short-lived and quickly disappeared, only to reappear later, in a slightly less advanced form, but more stable (week 47, relative F_2 score=0.868).

For both words, we also selected the "best" recordings of each week, that is, the recordings in which the labels and isolated vowels were the most recognizable and reliable. The resulting development curves follow the same pattern as described in Figures 4 and 5 (Figure 7).

Figure 7. Athena's mean F_2 of the vowel [ʊ] as in "wool" and "wood" in reference to her trainer's mean F_2 – Best weekly recordings

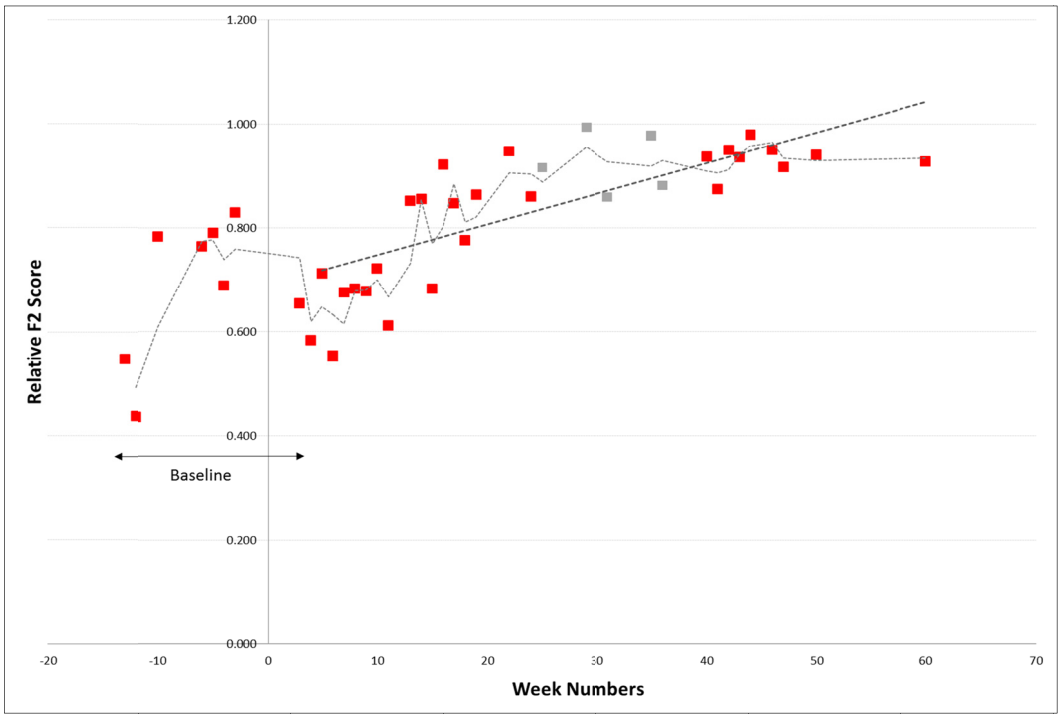


A. Wood B. Wool

These graphs represent the frequency values of Athena's F_2 against her trainer's for the vowel [ʊ] as contained in "wood" (A) and "wool" (B). Outliers have been removed according to the modified Thompson tau technique. Data points in grey represent the period of data loss/poor recordings. Only the "best" weekly recordings were incorporated for acoustic analysis.

Production of "paper". Unlike the two precedent labels whose vocal development took the shapes of multiple "U's", "paper" followed a more linear learning process (Figure 8).

Figure 8. Athena's mean F_2 of the vowel [ə] as in "paper" in reference to her trainer's mean F_2 (10/3/2013–3/8/2015)

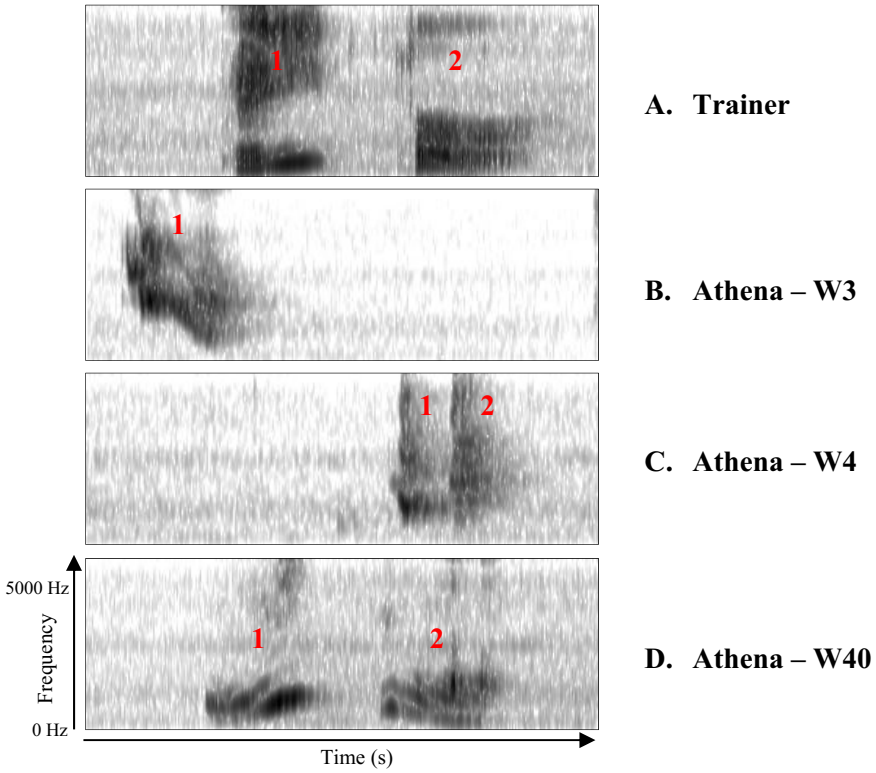


This graph represents the frequency values of Athena's F_2 against her trainer's for the vowel [ə] as contained in "paper". Outliers have been removed according to the modified Thompson tau technique. Data points in grey represent the period of data loss/poor recordings.

Athena's utterances of "paper" and productions of the vowel [ə] improved gradually from week 6 (relative F_2 score=0.554) until week 29 when they were spoken in a quite accurate fashion, scoring a maximum relative F_2 value of 0.995. Her ability to vocalize [ə] reliably further stabilized within a relative F_2 range of 0.882–0.980. At a very early stage (week 4), we could already discern a two-syllable sound when Athena was shown a piece of paper (Figure 9C). It sounded more like parrot noise than speech, but it improved steadily, resembling "pa-per" in melody (acoustic envelope). She was able to preserve this two-syllable structure and rarely reversed to uttering only one syllable when labeling

paper (Figure 9B). Although she had certainly made good progress, successfully shaping the vocalizations from "ay-ah", to "ay-er", Athena hadn't been able yet to produce a clear and decent "paper" by week 40 (Figure 9D).

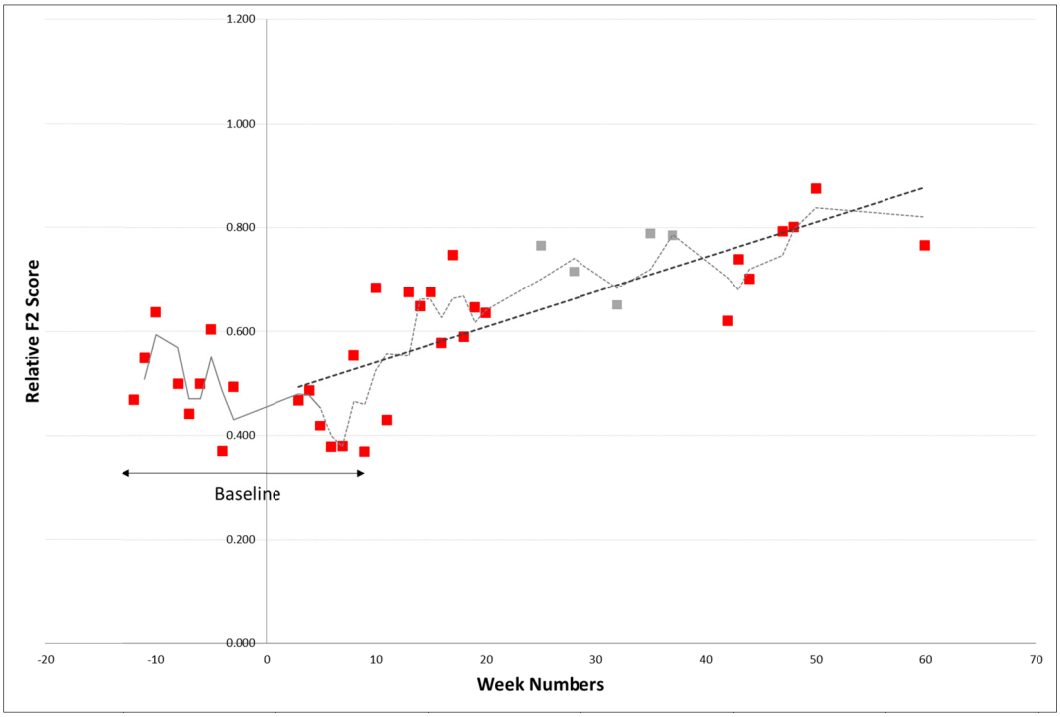
Figure 9. Spectral comparison of the speech utterance "paper"



These four wideband spectrograms illustrate the label "paper" of the trainer (A) compared to Athena (B, C, D). Spectrogram B was computed on 1/17/14 (week 3), spectrogram C on 1/23/14 (week 4) and spectrogram D on 10/4/14 (week 40). Two syllables (1 & 2) are clearly visible in A, C and D, while only one syllable appears in B.

Production of "truck". The development curve that characterizes "truck" is very much like "paper" (Figure 10).

Figure 10. Athena's mean F_2 of the vowel [ʌ] as in "truck" in reference to her trainer's mean F_2 (10/7/2013–3/8/2015)

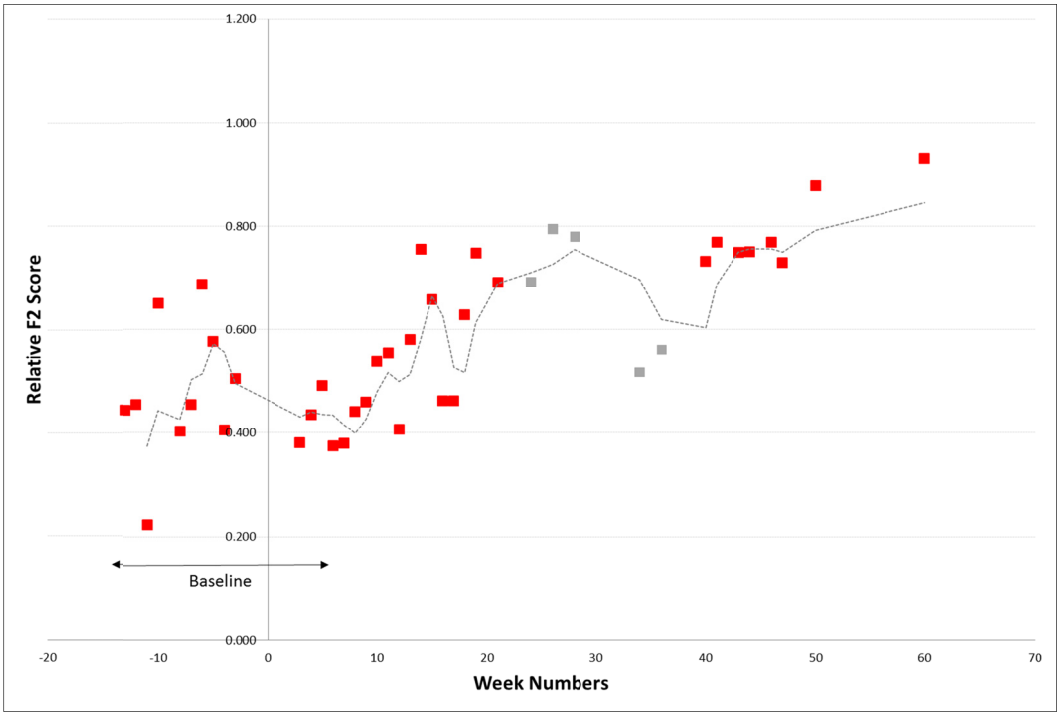


This graph represents the frequency values of Athena's F_2 against her trainer's for the vowel [ʌ] as contained in "truck". Outliers have been removed according to the modified Thompson tau technique. Data points in grey represent the period of data loss/poor recordings.

The acquisition pattern is also gradual, although the pace was somewhat more accelerated for "truck" at the beginning of the year between weeks 7–17 as the relative F_2 score increased by 0.367 points vs. 0.171 for "paper". Another notable difference is the sudden drop in accurate pronunciation of the vowel-like sound [ʌ] contained in "truck" at week 37 (relative F_2 score collapsed from 0.786 to 0.620 in one and a half month). This loss of accuracy was brief as Athena's correct production reappeared about six weeks later. A similar developmental pattern was also observed in "wood" and "wool".

Production of "nylon". The course of development of "nylon" appeared to follow the characteristic pattern of the early words "wood" and "wool" (Figure 11).

Figure 11. Athena's mean F_2 of the vowel [ɑ] as in "nylon" in reference to her trainer's mean F_2 (10/2/2013–3/8/2015)



This graph represents the frequency values of Athena's F_2 against her trainer's for the vowel [ɑ] as contained in "nylon". Outliers have been removed according to the modified Thompson tau technique. Data points in grey represent the period of data loss/poor recordings.

However, the "explosion" observed in "wood" and "wool" during the early learning phase was less sudden and less rapid with "nylon", although still considerable. "Nylon" also differed from the other two labels with regard to the maximum relative F_2 score that was reached after the initial growth phase (0.755 for "nylon" at week 14, vs. 0.871 for "wool" at week 13 and 0.994 for "wood" at week 10). Without a stabilization period, the

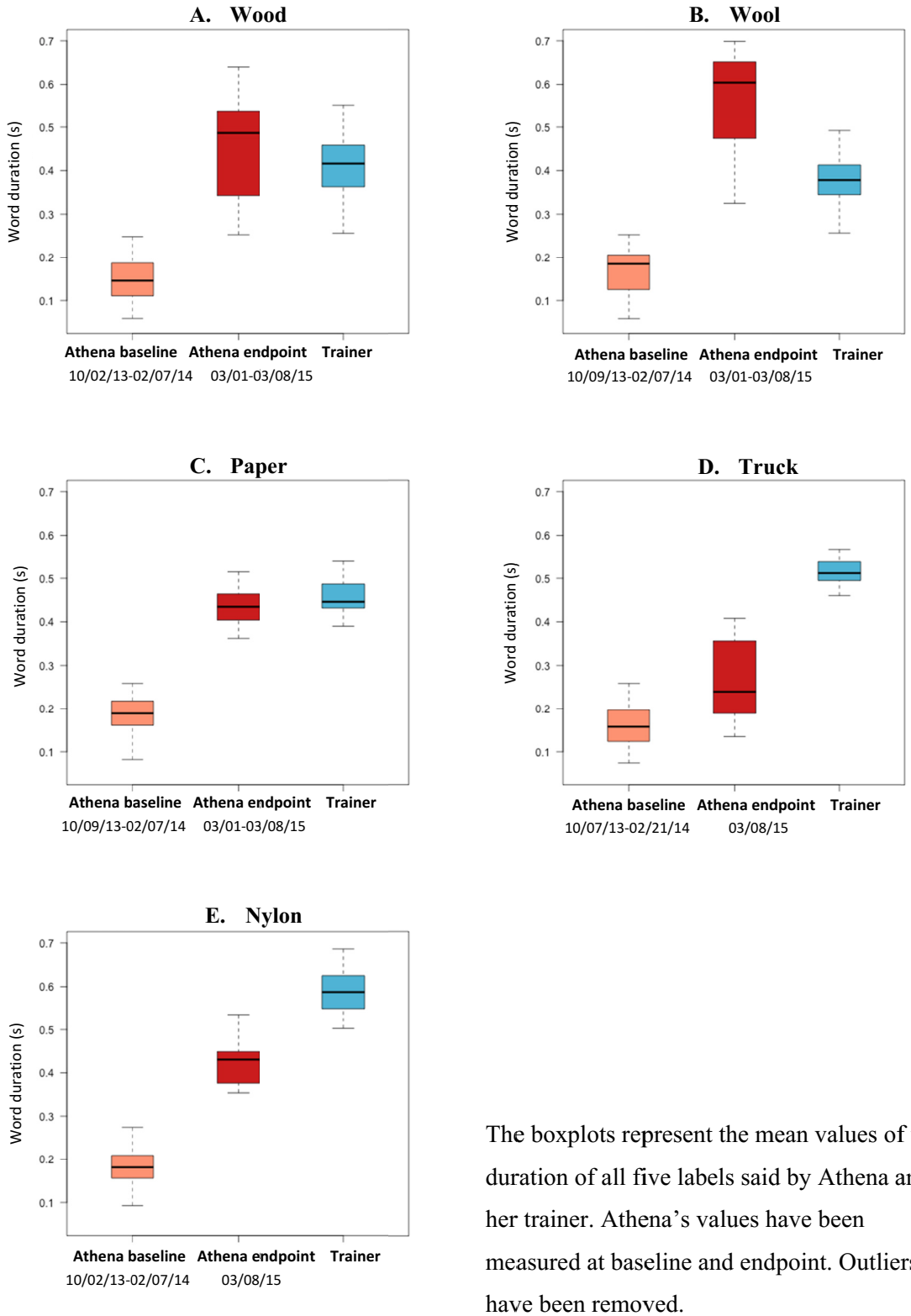
curve decreased abruptly around week 16 (relative F_2 score=0.462) but increased suddenly again, then plateaued over the summer, never exceeding a relative F_2 score of 0.794. As seen in the analysis of "wood" and "wool", the results show another drop right before the start of the Fall semester, with a subsequent recovery around week 40 (week 40 for "wool" and "nylon", week 42 for "wood"). At that time, Athena's pronunciations of the vowel [ɑ] and of the labels became suddenly more recognizable and reliable. While her vocalizations had remained below a maximum F_2 relative score of 0.794 since the beginning of the year, she made a breakthrough at the end of 2014 when she scored 0.877, a path she continued in 2015 (F_2 relative score of 0.929 in March 2015).

Word Duration Analysis

Comparisons of means of duration for each word revealed that Athena's productions in the beginning of the second year of training (at endpoint) were significantly longer than those at the beginning of the first year of training (at baseline) (two-tailed t test; "wood", p= 7.11E-9; "wool", p= 1.71E-05; "paper", p= 1.79E-06; "truck", p= 0.008; "nylon", p= 1.26E-7) (Figure 12).

Athena's sound repertoire at baseline was composed merely of calls emitted in short bursts, whistles and squeaks, with a mean duration averaging 0.168s. As her attempts and efforts to learn words and imitate speech began, longer vowel-like and speech-like sounds began to replace her vocalizations (overall mean duration across all labels at endpoint: 0.426s, a two-and-a-half-fold increase from baseline).

Figure 12. Boxplots of the mean word duration across all labels said by Athena and her trainer



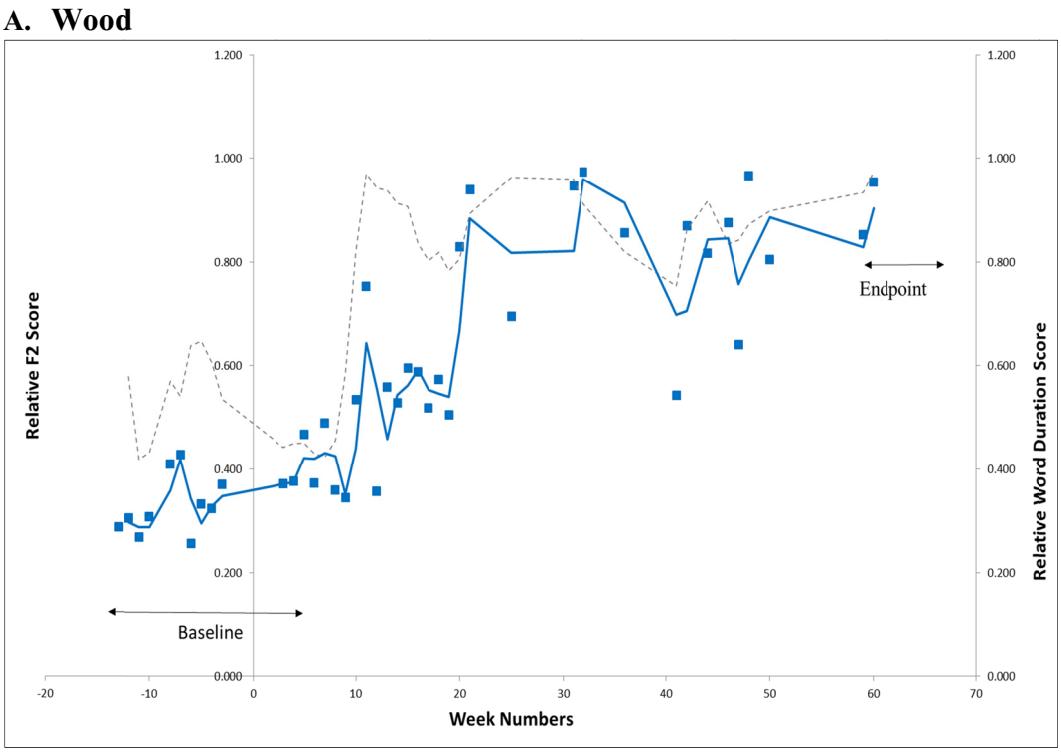
The boxplots represent the mean values of the duration of all five labels said by Athena and her trainer. Athena's values have been measured at baseline and endpoint. Outliers have been removed.

Although “nylon” and “paper” are the only two two-syllable words of the repertoire, they did not register the highest increase in absolute mean value between baseline and endpoint (0.244s and 0.237s, respectively) whereas the single-syllable words “wool” and wood” did (0.387s and 0.302s, respectively) (Figures 12A, 12B, 12C and 12E). These surprising observations most likely resulted from the strategy of Athena’s trainer to use stressed syllables to draw her attention (e.g. “wood-de”, “wo-ol”). For both tokens, given we used separate recordings for the trainer’s values, his mean durations were lower than those of Athena. “Paper” recorded the closest match to the trainer’s mean value at the endpoint across all labels, with 0.430s and 0.455s for Athena and her trainer, respectively (Figure 12C). Another notable result was that Athena seemed to display more variability at the endpoint, across all labels. Moreover, her variability was even higher on one than two-syllable words (standard deviations: “wool”=0.130, “wood”=0.121,”truck”=0.088, “paper”=0.075, “nylon”=0.054).

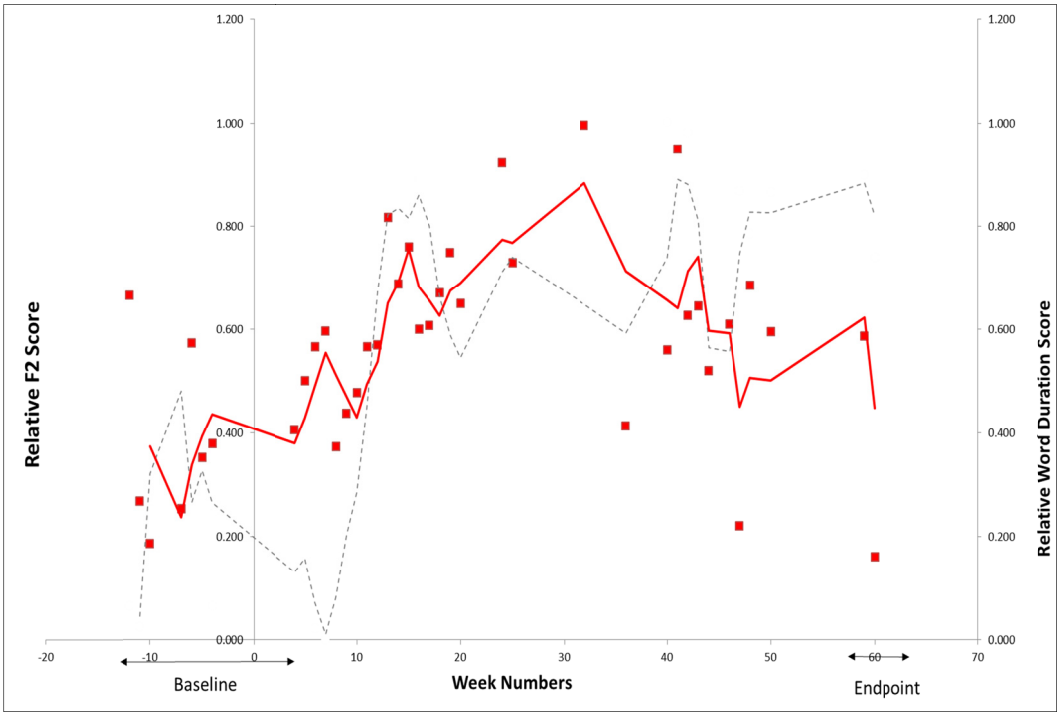
When we superimposed Athena’s relative duration value lines over the trajectories of her relative F_2 scores, we obtained a rather interesting image: the same general pattern of changes that affected Athena’s F_2 values appeared to also affect the length of her labels (Figure 13). However, more remarkably, until about week 10, both curves evolved symmetrically in the opposite direction, as if one mirrored the other. After that period, all productions experienced growths in duration in the same way that Athena’s second formant values increased (except around week 19 for “truck” and week 40 for “paper”, when the relative word duration value curves became erratic and no longer followed the F_2 trajectories). Finally, beginning at week 47, all labels became shorter while the F_2 relative scores increased. Unfortunately, we do not have enough data

points to confirm if this is another "mirror effect". The trend was especially severe with "wood" and "wool" whose durations were reduced by 40.7% and 21.1% between weeks 47–50, respectively. However, the last recordings made on March 8, 2015 reversed the trend, except for "truck" and "nylon". "Wood" increased in length by about 30%, while "paper" was augmented by twice as much. "Wool" was inconclusive.

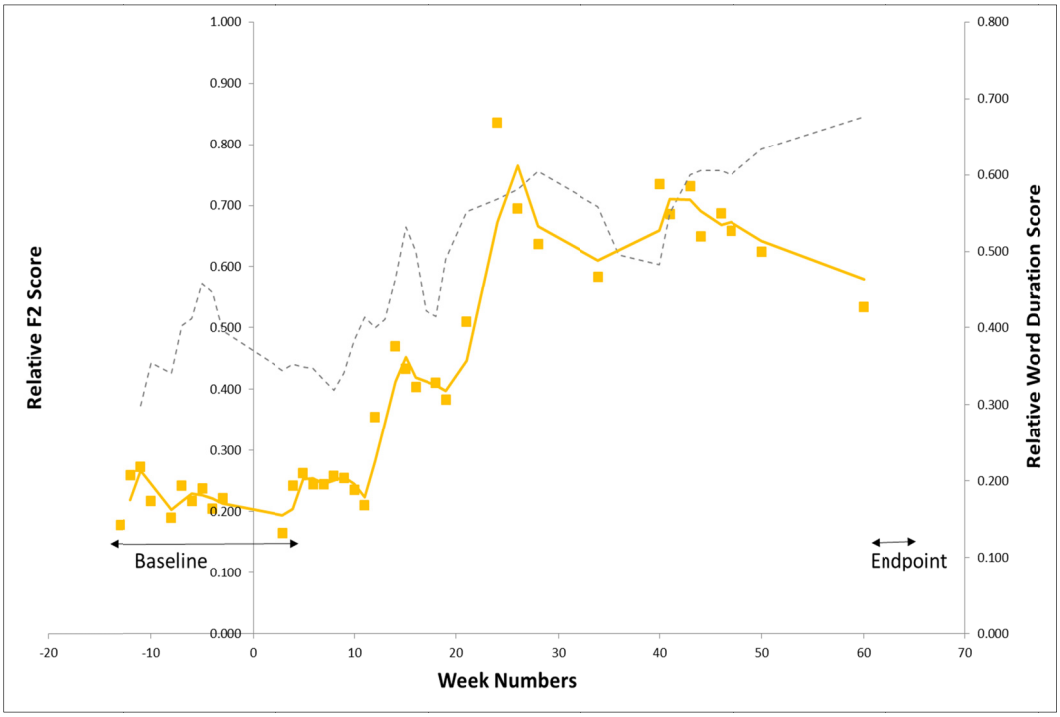
Figure 13. Athena's mean word duration curve and F_2 trajectories across all labels



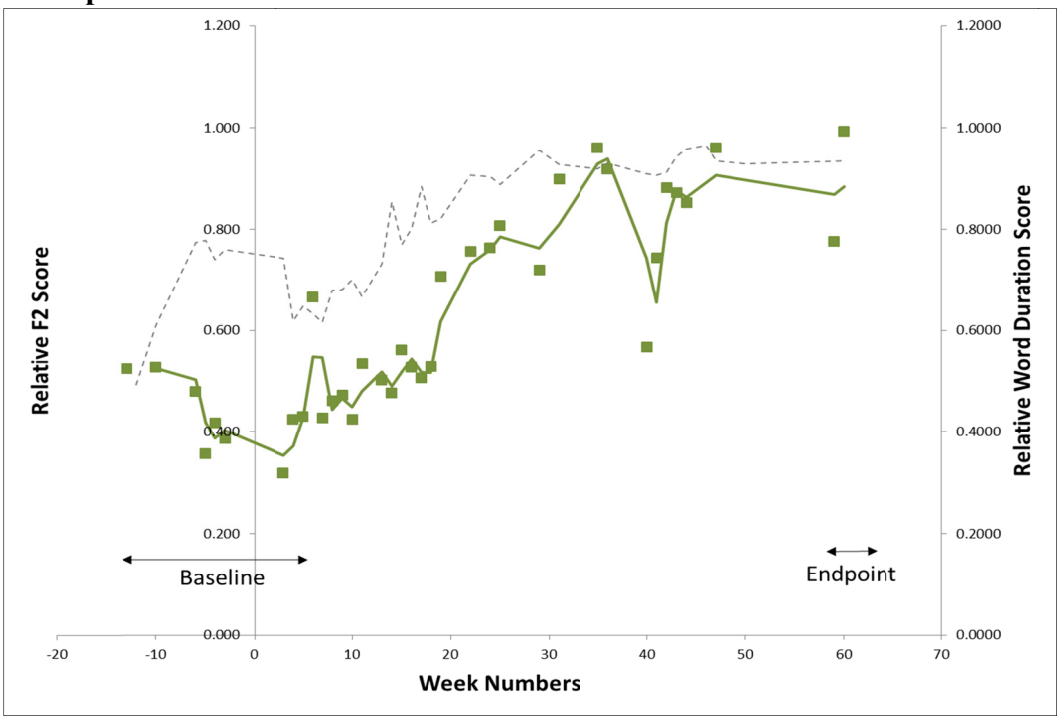
B. Wool



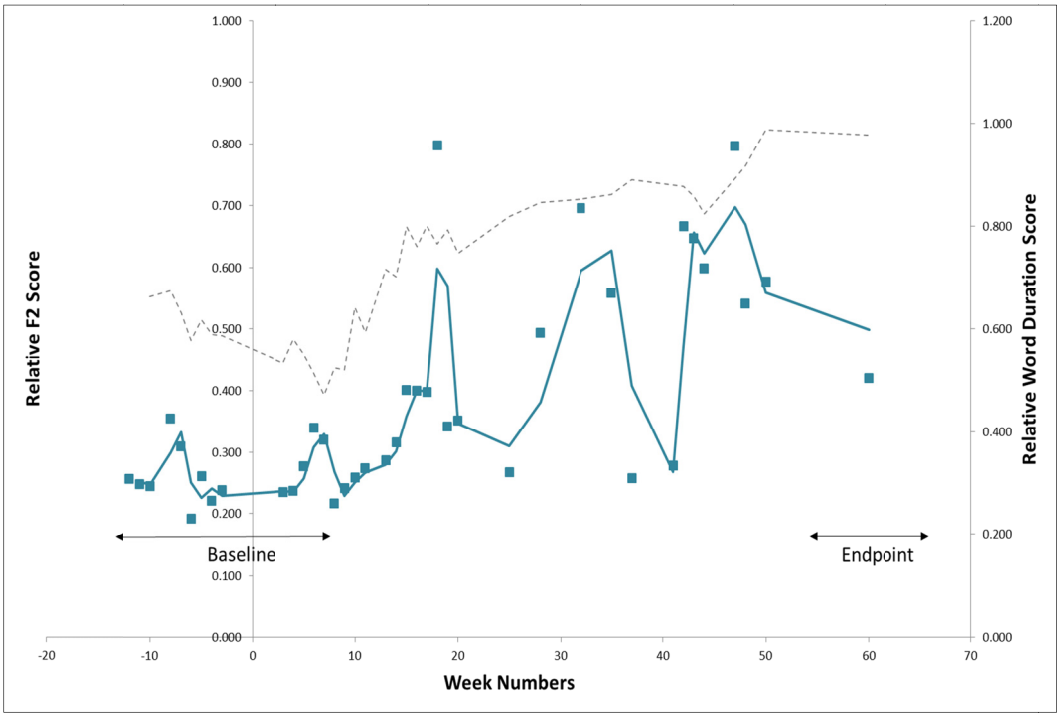
C. Nylon



D. Paper



E. Truck



The above figures represent the relative word duration scores of each label (A. Wood, B. Wool, C. Nylon, D. Paper, E. Truck), from baseline through endpoint. The dashed, grey line is the curve for F_2. Outliers have been removed according to the modified Thompson tau technique.

Athena's F_2 Formant Matching with Griffin - Acoustic Analysis per Label

Griffin, Athena's lab companion, participated in about 5% of the training sessions, acting as a model for her as she began to learn new vocal labels. Given such a limited exposure, we expected the influence of Griffin's pronunciation on Athena's vocal development to be minimal. Yet, since Griffin was in Athena's presence at all times (except for recording sessions), we could not avoid the fact that he may somehow have influenced Athena's learning patterns and pronunciation. King et al. (2005) reported that female cowbirds (*Molothrus ater*) can shape the vocal development of young males, even though females lack the ability to sing. When housed with males in pairs or trios, females use social cues including wing strokes and gapes as positive feedback to retain specific behaviors associated with song development. Therefore, to determine if Athena could have been affected by Griffin social behavior, we compared in the following section Athena's second formant with her parrot companion's corresponding formant. Mean formant values for Griffin across vowels are provided in Table 5.

Table 5. Means of F_2 for Griffin, an African Grey parrot, across vowels

	[ʊ] as in w<u>oo</u>l, w<u>oo</u>d		**[ə] as in pap<u>e</u>r**		**[ʌ] as in tr<u>u</u>ck**	
	F_2	**Nb. of samples**	**F_2**	**Nb. of samples**	**F_2**	**Nb. of samples**
Griffin	941[a] (± 65)	47 (23 "wool" + 24 "wood")	1681 (± 73)	20	1227 (± 45)	21

[a] F_2 is the mean of 938 ("wool") and 943 ("wood")

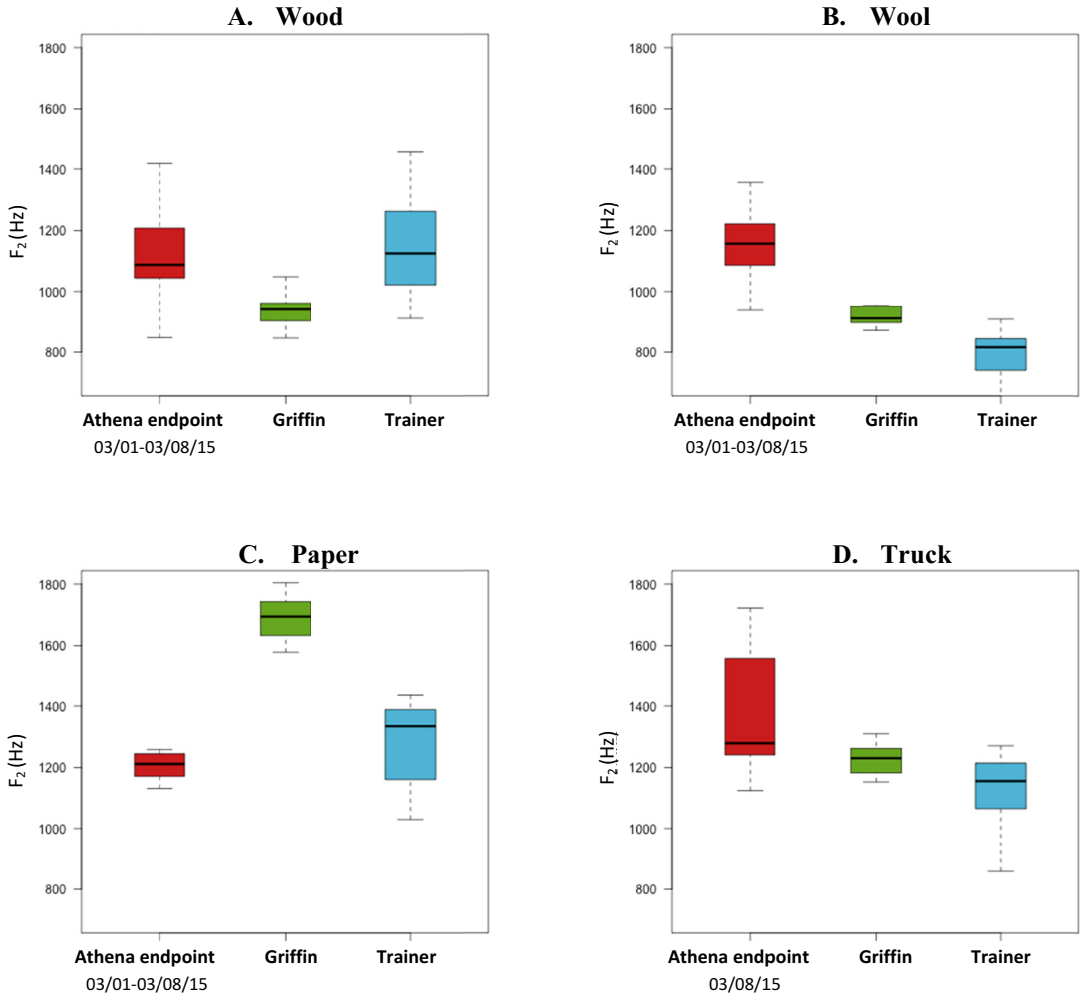
Notes:

- F_2 for "nylon" is not available as this word is not part of Griffin's vocabulary
- Outliers have been removed according to the modified Thompson tau technique.

This table lists the mean formants values of Athena's lab companion Griffin that were used in the vocal development comparative analysis. Standard deviations are listed in parentheses.

As expected, Athena's second formant frequencies of the vowels [ʊ], [ə], and [ʌ] differed from the values obtained when recording Griffin. A two-tailed t tests across all four tokens that contained these vowel-like sounds showed that Athena's mean F_2 frequencies were significantly different from her lab companion's corresponding data ("wood", p= 1.96E-04; "wool", p= 0.004; "paper", p= 2.49E-15; "truck", p= 0.030;). The boxplots in Figure 14 confirm this statement.

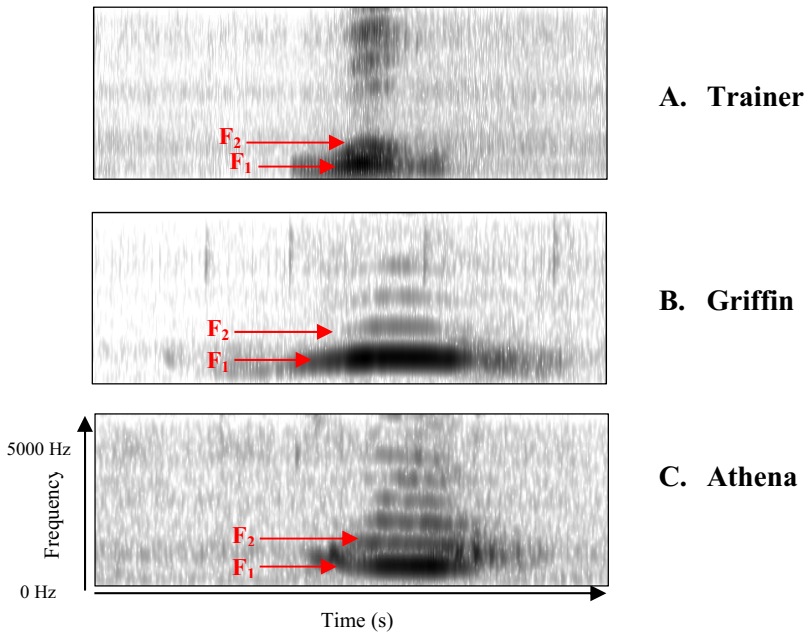
Figure 14. Boxplots of the mean F_2 across four labels of Athena, Griffin and Trainer



The boxplots represent the mean values of the duration of four labels that contained the vowel-like sound [ʊ], [ə], and [ʌ] said by Athena, Griffin, and the human trainer. Athena's values have been measured at endpoint (March 2015). Outliers have been removed.

Notably, her mean F_2 frequencies were higher than her lab companion's corresponding formant values for the vowels [ʌ] and [ʊ] (but not for [ə]). In the case of [ʌ] like in "truck", the boxplot showcases an overlap of Athena's and Griffin's F_2 frequencies and close mean values (Athena=1393, Griffin=1227). Despite significant differences in means for [ʊ], the spectrograms exemplifying the utterance "wool" of Athena and Griffin share similarities in shape (Athena's mean F_2=1135, Griffin's mean F_2=938; Figure 15).

Figure 15. Spectral comparisons of the speech utterance "wool"



This figure represents three wideband spectrograms illustrating Athena's trainer (A), Athena's laboratory companion Griffin (B) and Athena (C) saying the label "wool". The first two formants, F_1 and F_2, are indicated. Note the similar shapes of the parrots' spectrograms, and also the resemblance with the trainer's. Athena's spectrogram is dated from 3/25/14 (week 13).

It is also of interest to note that Athena's overall values were a closer match to her human model than to her conspecific's. Her range of frequency values for F_2 covers only 36.9% of Griffin's but 52.5% of her trainer's.

Chapter 4

Discussion

How children acquire words is one of the central themes of language research. The parallels that have been drawn between birdsong and speech learning have given scientists the opportunity to use birds as models for the study of vocal learning, but the lack of referentiality in song limits the utility of avian models. By testing the rules for parrots and unveiling some of the mechanisms underlying their ability to acquire speech-like sounds that can be used referentially, researchers might be able to make predictions about how children learn their first words. Yet, very little is known about early vocal development of parrots. For that purpose, we examined the patterns and timing of vocal development of a juvenile African Grey parrot who is learning referential English. In particular, we studied the acoustic developmental pattern of vowel-sounds contained in selected labels the bird attempted to acquire over the course of fifteen months. We predicted that as young children occasionally do, our subject Athena would, in some cases, provide evidence for a regression in her phonology: she would produce an early word in a near perfect form, and then show regression in phonetic accuracy before producing the label correctly again at a later stage.

Hypothesis Tested: Phonological Regression in a Young African Grey Parrot Who Is Learning Referential English

Our hypothesis was proven at least partially correct. The isolated labels “wood” and “wool” were pronounced in a relatively accurate form at a very early stage while

other words were still only amorphous sounds. Then, Athena's pronunciation of these tokens became worse and reverted back to unstructured sounds, but eventually improved again a few weeks later, following a U-shaped curve. These findings are consistent with the phenomenon of phonological regression observed in children and provide strong support to the claim that Athena may have also experienced such a non-monotonic pattern of development. However, unlike children who generally experience one single instance of U-shaped development, her course of developmental learning was interspersed by at least one[1] additional U-shaped curve. A possible explanation for these later results may come from the factors contributing to phonological regression, and this led us to wonder if, in general, phonological regression comes from the same sources in parrots and humans. In humans, such a nonlinear trajectory is generally associated with the formation of rules and the emergence of organization and systematic strategies for pronunciation (Bleile & Tomblin, 1991; de Villiers & de Villiers, 1979; Leopold, 1947; Vihman, 2014). Based on these findings, we can suggest that, shortly after the baseline period, Athena might have transitioned from an "associative" to a "rule-based" behavior much like a child would have shifted from a pre-rule stage to an adult rule-based system (Rogers, Rakison, & McClelland, 2004). According to this explanation, the early, accurate utterances of "wood" and "wool" were merely mimetic forms that were rapidly overcome by more systematic approaches required for the acquisition of referential English. Therefore, consistent with the explanation of phonological regression in children, the adoption by Athena of a rule and template-based system resulted

[1] The plots showed multiple additional U-shaped curves, but the poor quality of the data between 06/17/14 – 09/30/14 made it difficult to extract reliable and valid conclusions, and therefore we focused in the discussion on the recording periods that generated datasets of high quality.

temporarily in a performance decline that was illustrated by a drop within the first U-shaped curve (week 17).

Concerning the occurrence of recurring U-shaped patterns—for example, the curve observed at the beginning of the last semester (week 42)—potential contributing factors may include training and social interaction. Notably, Athena's pronunciation of vowels improved at the same time the students were returning from the summer break. Indeed, with the lab fully staffed, Athena was submitted to a more regular training schedule and also benefited from a greater variety of exercises. The presence of additional students also meant more people with whom she could interact, not only vocally but also socially. Home-raised African Grey parrots are known for establishing strong social contacts with their caretakers and Athena indeed strongly bonded with some of the students, enjoying particularly standing on their shoulders or being tickled. Taken together, her increased exposure to training and enhanced social enrichment enabled her to fully exploit her abilities and to learn labels more readily (Pepperberg, 1994, 2007). In contrast, the negative effects of the winter recess were reflected in the drop in her level of vocal development in January (week 5) after she had failed to maintain the initial improvement in performance she had acquired during the baseline period. Another possible explanation for the occurrence of an alternating pattern of multiple U-shaped curves comes from cognitive overload. At the time of the experiment, Athena was not solely trained on the five labels that were the focus of the project, but rather on a full repertoire of sixteen words. In addition, students were trying to teach her new words, outside of the standard training sessions, by connecting labels to objects. Therefore, the execution of new strategies for maintaining certain labels and an increase in vocabulary

size might have led to an overload of her cognitive system. Possibly this overload resulted in a temporary decrease in processing efficiency, explaining the decrement within the U-shaped pattern. However, Athena may then have shifted more cognitive resources to the task, because the processing capabilities were rapidly recovered and the performances improved.

Cognitive overload might also be responsible for the phenomenon documented at baseline when the curves for the relative duration and F_2 scores evolved symmetrically in opposing directions, as if one was mirroring the other (Figure 13). This pattern reflects Athena's failure to process both factors together in her attempts to match her human model: reduce F_2 frequencies to achieve the target formant values for vowels and augment utterance duration to replace calls and whistles by speech-like vocalizations. As with phonological regression, we can posit that here also, the demands exceeded the available resources and temporarily prevented her cognitive system from processing those two high-level tasks simultaneously, thus resulting in developmental tradeoffs between F_2 and duration accuracy.

We can also speculate that the existence of multiple U-shaped curves is inherent to Athena's learning process for early labels. In other words, while human infants seem to require only one regression period, parrots might experience two or multiple regression cycles during their normal learning process. It is possible that without such a repeated pattern of alternating U-shaped curves, she would not be able to fully acquire the specific labels on which she was being trained. Moreover, children are learning species-specific vocalizations, and Athena was learning heterospecific vocalizations; maybe heterospecific learning requires extra processing power and multiple stages of regression.

Comparing Developmental Patterns of Word Learning between a Young African Grey Parrot, Human Children and Other Parrots

The results obtained in our current study clearly point out to potential similarities between children and a parrot regarding the phenomenon of phonological regression and the sources from which it may arise. Moreover, the results also provided a broader picture of the dynamic process of word learning in parrots. In the following section we compare some of the characteristic developmental learning patterns among Athena and human children and other parrots.

Trend analyses of the course of vocal development of Athena reveal a gradual lowering in mean F_2 as a function of increasing age. The F_2 value for baseline and endpoint averaged 1680 Hz and 1184 Hz, respectively. These results are consistent with previous developmental studies of vocal development in young infants and children (Gilbert, Robb, & Chen, 1997; McGowan, McGowan, & Denny, 2014). Gilbert et al. reported a 24.2% decrease in the second formant frequency across the 15- to 36-month age period of four children, from 2558 Hz to 1938 Hz. Athena's range of frequency for F_2 was compressed at baseline ([1637 Hz–1700 Hz]) but it decreased with age (data at endpoint: [1124 Hz–1393 Hz]). These findings are in agreement with the classic study of Kent & Murray (1982) that recorded vocalizations of infants at 3 and 9 months of age and documented an increase in the F_2 range of 700 Hz between the two surveys.

Interestingly, across Athena's baseline period, the data show no significant difference between the mean F_2's of any of the vowels, suggesting an overlap in her vowel spaces. In humans, small vowel area is considered an indicator of unintelligible speech (Ferguson & Kewley-Port, 2007). Athena's mispronunciations and tightly

clustered vowel plot at that time (Figure 2) confirm that this robust relationship between vowel space and speech can be applied to parrots. However, as her overall pronunciation improved, her vowels became increasingly differentiated (Figure 2). What could account for the observed separation of vowels? Anatomical changes could be responsible. Indeed, a child's vocal tract undergoes dramatic changes from infancy to adulthood that influence developmental changes in formant frequencies and contribute to the development of vocalizations (reviewed in Mugitani & Hiroya, 2012). As we observed shifts in the frequency of Athena's F_2, the maturation of her vocal tract might have played a role in the dispersion of her vowels. Sadly, we cannot draw any conclusion from these observations because to date, no study has documented the existence of a critical period for reorganization of the vocal tract in young parrots. Another possible explanation for the observed dispersion of Athena's vowels comes from the relationship between learning and vocal perception and production. As seen earlier, songbirds, like humans, learn the vocal sounds from adult "tutors" during an early phase of learning that is primarily perceptual. Then, they use auditory feedback to gradually form their own song through a sensorimotor process of matching their own vocal output to the memory of the tutor sounds (reviewed in Doupe & Kuhl, 1999). Parrots, unlike songbirds, do not learn "songs" but rather learn complex calls (e.g. "begging calls", "contact calls") from their parents (Berg, Beissinger, & Bradbury, 2013). Therefore, it is possible that Athena might have used the labels spoken by her trainer to guide her own vocal production of the vowels. If this is the case, she might have used the perceptual representations of the vowels stored in her memory as targets to match when producing her vocalizations. By comparing her developing vocalizations with these "templates", she eventually might

have converged towards the target vowels of her trainer. In support of this explanation is the vowel distribution map in Figure 2 that outlines the trend towards expanding the vowel space and matching her trainer's values. This account is also supplemented by the results presented in Figure 3 that show the parrot's effort towards creating an accurate imitation of the formant frequencies of her trainer. Despite the apparent separation of the vowels at the endpoint, they were still relatively clustered compared to her trainer. This relative clustering could be attributed to the fact that even in the defined endpoint of the study, Athena was still in the process of learning and therefore had not yet quite reached the full spectrum of F_2 vowel frequencies. Surprisingly, Athena had F_2 values at endpoint that were a closer match to her human model than to a conspecific (Griffin) (Figure 14). Vehrencamp et al. (2003) conducted playback studies of geographic dialects from wild parrot populations of orange fronted conures (*Aratinga canicularis*) and have demonstrated that birds reacted more strongly to "local" stimuli. That is, there must be filtering mechanisms that predispose the parrots to attend to these specific regional signals within the environment. Therefore, Athena might have an innate focus on sounds that are species-typical, not human. It is however possible that in the current study, timing and the amount of exposure to speech are factors that might have influenced her utterances as are often described in children and in songbirds (Kuhl & Meltzoff, 1996; Kuhl et al., 2005). Because the modeling sessions with Griffin started late in the learning process and there were only few of them, Athena might not have been particularly sensitive to his input. In contrast, because she was spending a great deal of time with her principal trainer since she was five months old, her vocalizations might have been greatly

influenced by him. Finally, the observed U-shaped pattern in F_2 frequency during label acquisition also provides significant support to this explanation.

As described in the previous section, “wood” and “wool” followed a U-shaped curve for phonological development, with largely accurate earliest forms and subsequent distorted productions. It is interesting to note that the exact rendition of “wool”, which occurred three weeks after the correct vocalization of “wood”, might have been facilitated by “wood” because of the phonetic similarity. However, the other words in process of acquisition, which differed significantly from the labels already in Athena’s repertoire, did not show any evidence of early accuracy but rather followed a gradual increase of relative F_2 score, reflecting a more straightforward process of improvement. Nevertheless, the two syllables of “paper” could be heard as early as week 4 when Athena was shown a piece of paper (Figure 9C). She lacked the accuracy of the vowel pronunciation, hence the speech clarity, but her vocalizations resembled “pa-per” in rhythm (acoustic envelope). The lack of lips makes it very difficult for a parrot to render the sound “puh”, and this explains in part why Athena had not be able to utter a clear and decent “paper” by week 40 (Figure 9D). To produce such plosives, Grey parrots seem to need to learn to use esophageal speech (Patterson & Pepperberg, 1998). However, she maintained a “vocal contour” of the word, and rarely reversed to uttering only one syllable, improving her pronunciation steadily, first saying “ay-ah” then “ay-er”. This strategy was also one of the forms of vocal learning adopted by Alex, the famous parrot. His words were unstructured when they emerged for the first time, shaped only by the acoustic envelope, then with the vowels and finally the consonants (Pepperberg, 1999).

Surprisingly, Athena seemed to rarely engage in private vocal practice. We expected that by "talking" to herself privately, she would consolidate her knowledge and accelerate her acquisition of labels. But the monologue samples obtained from the brief intervals of time she was left alone by her trainer during the recording sessions revealed that, unlike babies who actively babble alone and experiment sounds in their cribs, or even Alex who did practice in private to acquire labels such as "none" (Weir, 1962; Pepperberg, Brese, & Harris, 1991), she did not vocalize. Even "wool", which was the most likely candidate for practice as it only differs from the already acquired "wood" by one consonant, did not appear in the recordings. A possible reason to the absence of monologue was that Athena was less motivated than Alex and did not attempt to practice outside the boundaries of the sessions. Alternatively, it is possible she might have instead engaged in covert speech in the form of mental play as some children often do (Kuczaj & Bean, 1982). Another possible reason is that she might have practiced vocalizations at night, when we were not recording.

During the course of our study, it also became evident that Athena, like Alex, occasionally showed a lack of motivation and would not engage in a task if she was not interested. In the absence of food reward, we had to get her attention with objects about which she was curious or with which she liked to play. Unfortunately, since she had not learned yet to use "want" to choose her objects like Alex did, we had to keep the training sessions brief due to her short attention span (Pepperberg 1999).

Limitations and Future Directions

On a final note, we acknowledge that the results of the present study should be interpreted with caution. Limited pool subject studies are often criticized. The small sample size and the limited number of speech labels may reduce the likelihood that we are observing a real effect and that the results we obtained are reproducible. However, according to Triana and Pasnak (1981), a "power study" with a single or a few subjects has value because the ability showed by one individual is within the scope of the entire species. Furthermore, the first investigations on child language which have provided valuable data to the field of language acquisition focused on the intensive case study of only one subject at a time. They were often diaries tracking the linguistic development of a single child based on the parents' observations (de Villiers & de Villiers, 1979; Leopold, 1939, 1947).

A direct consequence of the limited subject pool is that the fate of the entire experiment (i.e. generation of data) depended on Athena's motivation to cooperate and vocalize.

Despite these issues, we note that the parallels between a parrot learning to use referential English and a child learning to speak English are striking. From analogous brain regions to vocal learning and social influences, both systems share commonalities. However, differences between human and parrot speech certainly exist, one being the ability to convey abstract thought and semantic complexity (Berwick et al., 2011). Therefore, we should take a cautious approach when extrapolating any conclusion from our results as to how humans may have developed the ability to speak.

These limitations point to future lines of research that would use a larger sample, and a longer time span to describe longitudinal vocal development in an African Grey parrot. Because U-shaped developmental curves have been observed in a wide variety of other learning and cognitive processes, it would be particularly interesting to test a juvenile Grey parrot with tasks that involve U-shaped behavioral patterns in humans (Gershkoff-Stowe & Thelen, 2004). If the results show that the bird develops these other skills according to the same trajectory, "do well, then do worse before doing better again", then this new evidence would compete with the traditional monotonic and cumulative model of improvement with time. Further research would be then required to narrow the cognitive and learning abilities that fit this novel nonlinear learning pattern.

Conclusion

Overall, we have shown that a juvenile African Grey parrot who is learning referential English, shares with children who are learning to speak, several developmental patterns. Our results demonstrate that parrots, like human infants, can pass through period(s) of phonological regression in accuracy in which an early set of words follow a nonlinear pattern of development. According to this model, a child's or a parrot's first words may be produced in a way that exceeds their speech ability at the time, then deteriorate, only to revert back, later, to their correct forms. Although regressions can arise from a variety of sources, some of them, in particular the emergence of organization and rule-based speech patterns, might also hold true for parrots. Acquiring a code for referential communication requires the ability to not only connect labels with objects but also to adhere to rule-based strategies to enable further learning and speech development.

References

Anisfeld, M. (1984). Language development from birth to three. Hillsdale, NJ: Lawrence Erlbaum Associates.

Baptista, L.F. (1983). Song learning. In A.H. Brush & G.A. Clark Jr (Eds), *Perspectives in ornithology* (pp 500-506). Cambridge: Cambridge University Press.

Beckers, G.J.L., Nelson B.S., & Suthers, R.A. (2004) Vocal-tract filtering by lingual articulation in a parrot. *Current Biology,* 14, 1592–1597.

Berg, K.S., Beissinger, S.R., & Bradbury, J.W. (2013). Factors shaping the ontogeny of vocal signals in a wild parrot. *Journal of Experimental Biology*, 216, 338–345.

Berwick, R.C., Okanoya, K., Beckers, G.J.L., & Bolhuis, J.J. (2011). Songs to syntax: the linguistics of birdsong. *Trends in Cognitive Sciences,* 15, 113–121.

Bleile, K., & Tomblin, K. (1991). Regressions in the phonological development of two children. *Journal of Psychological Research,* 20(6), 483–99.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, 5(9/10), 341-345.

Bolhuis, J.J., & Everaert, M. (2013). *Birdsong, speech, and language: Exploring the evolution of mind and brain*, Cambridge, MA: MIT Press.

Bolhuis, J.J., Okanoya, K., & Scharff C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Review Neuroscience,* 11, 747–759.

Bottoni, L., Masin, S., & Lenti Boero, D. (2009). Vowel-like sound structure in an African Grey Parrot (*Psittacus erithacus*) vocal production. *The Open Behavioural Science Journal,* 3, 53–68.

Burish, M.J., Kueh, H.Y., & Wang, S.S-H. (2004). Brain architecture and social complexity in modern and ancient birds. *Brain, Behavior and Evolution*, 63, 107–124.

Colbert-White, E.N., Covington, M.A., & Fragaszy, D.M. (2011). Social context influences the vocalizations of a home-raised African Grey parrot (*Psittacus erithacus erithacus*). *Journal of Comparative Psychology*, 125, 175–184.

Cruickshank, A.J., Gautier, J-P., & Chappuis C. (1993). Vocal mimicry in wild African Grey parrots *Psittacus erithacus*. *Ibis*, 135, 293–299.

Darwin C. (1871). *The descent of man, and selection in relation to sex*. Murray, London.

de Villiers, P.A., & de Villiers J.G. (1979). *Early language, the developing child series*, Cambridge, MA: Harvard University Press.

Doupe, A.J., & Kuhl, P.K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567–631.

Emery, N.J. (2006).Cognitive ornithology: the evolution of avian intelligence. *Philosophical Transactions of the Royal Society B*, *London. Biological Sciences,* 361, 23–43.

Ferguson, S. H., & Kewley-Port, D. (2007). Talker Differences in clear and conversational speech: Acoustic characteristics of vowels. *Journal of Speech, Language, and Hearing Research*, 50, 1241–1255.

Gilbert, H.R., Robb, M.P., & Chen, Y. (1997). Formant frequency development — 15 to 36 months. *Journal of Voice*, 11, 260–266.

Giret, N., Albert, A., Nagle, L., Kreutzer, M., & Bovet, D. (2011). Context-related vocalizations in African grey parrots (*Psittacus erithacus*). *Acta Ethologica,* 15(1), 39–46.

Johnson, W., & Reimers P. (2010). *Patterns in Child Phonology*. Edinburgh: Edinburgh University Press.

Kent, R.D., & Murray, A.D. (1982). Acoustic features of infant vocalic utterances. *Journal of the Acoustical Society of America*, 72, 353–365.

King, A. P., West, M J., & Goldstein, M H. (2005). Non-vocal shaping of avian song development: parallels to human speech development. *Ethology,* 111:101–107.

Klatt, D.H, & Stefanski, R.A. (1974). How does a mynah bird imitate speech? *Journal of the Acoustical Society of America,* 55, 822–832.

Konishi, M. (1965). The role of auditory feedback in the control of vocalizations in the white-crowned sparrow. *Zeitschrift für Tierpsychologie*, 22, 770–783.

Kroodsma, D.E. (1974). Song learning, dialects and dispersal in the Bewick's wren. *Zeitschrift für Tierpsychologie*, 35, 352–380.

Kuhl, P.K., Meltzoff, A.N. (1996). Infant vocalizations in response to speech: vocal imitation and developmental change. *Journal of the Acoustical Society of America*, 100, 2425–2438.

Kuhl, P.K, Conboy, B.T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: implications for the 'critical period'. *Language Learning and Development*, 1 (3–4), 237–264.

Kuczaj, S.A., & Bean, A. (1982). The development of non-communicative speech systems. In S.A. Kuczaj (Ed.), *Language development: Language, thought, and culture*. Hillsdale, New Jersey: Erlbaum.

Leopold, W. (1939). *Speech development of a bilingual child: a linguist's record. Volume I. Vocabulary growth in the first two years*. Evanston, IL: Northwestern University Press.

Leopold, W. (1947). *Speech development of a bilingual child: a linguist record. Volume II. Sound-learning in the first two years*. Evanston, IL: Northwestern University Press.

Lindblom, B. (1986). Phonetic universals in vowel systems. In J.J. Ohala & J.J. Jaeger (Eds), *Experimental phonology* (pp. 13–44). Orlando: Academic Press.

Marler, P. (1970b). A comparative approach to vocal learning: song development in white-crowned sparrows. *Journal of Comparative Physiological Psychology*, 71, 1–25.

Marler, P., & Slabbekoorn, H. (2004). *Nature's music: the science of birdsong*. Boston: Elsevier.

May, D. (2004). The vocal repertoire of Grey parrots (*Psittacus erithacus*) living in the Congo Basin. PhD thesis, University of Arizona.

McGowan, R. W., McGowan, R. S., Denny, M. & Nittrouer, S. (2014). A longitudinal study of very young children's vowel production. *Journal of Speech, Language, and Hearing Research,* 57, 1–15.

Mugitani, R., & Hiroya, S. (2012). Development of vocal tract and acoustic features in children. *Acoustical Science and Technology*, 33(4), 215–220.

Oller, D.K. (1986). Metaphonology and infant vocalizations. In B. Lindblom and R. Zetterstrom (Eds.), *Precursors of early speech* (pp. 21-36). Baskingstoke, Hampshire: Macmillan.

Patterson, D.K., & Pepperberg, I.M. (1994). A comparative study of human and parrot phonation: Acoustic and articulatory correlates of vowels. *Journal of the Acoustical Society of America*, 96, 634–648.

Patterson, D.K., & Pepperberg, I.M. (1998). A comparative study of human and Grey parrot phonation: Acoustic and articulatory correlates of stop consonants. *Journal of the Acoustical Society of America*, 103, 2197–2213.

Pepperberg, I.M. (1981). Functional vocalizations of an African grey parrot (*Psittacus erithacus*). *Zeitschrift für Tierpsychologie*, 55, 139–151.

Pepperberg, I.M., Brese, K.J., & Harris, B.J. (1991). Solitary sound play during acquisition of English vocalizations by an African Grey Parrot (*Psittacus erithacus*): Possible parallels with children's monologue speech. *Applied Psycholinguistics,* 12,151–178.

Pepperberg, I.M. (1994). Vocal learning in grey parrots (*Psittacus erithacus*): effects of social interaction, reference, and context. *Auk,* 111, 300–313.

Pepperberg, I.M. (1999) *The Alex studies: cognitive and communicative abilities of Grey parrots*. Cambridge: Harvard University Press.

Pepperberg, I.M. (2006). Cognitive and communicative abilities of grey parrots. *Applied Animal Behaviour Science,* 100, 77–86.

Pepperberg, I.M. (2007). Grey parrots do not always 'parrot': Phonological awareness and the creation of new labels from existing vocalizations. *Language Sciences*, 29, 1–13.

Pepperberg, I.M. & Carey, S. (2012). Grey parrot number acquisition: The inference of cardinal value from ordinal position on the numeral list. *Cognition*, 125, 219–232.

Peterson, G.E., & Barney, H.L. (1952). Control methods used in a study of the identification of vowels. *Journal of the Acoustical Society of America*, 24, 175–184.

Petkov, C.I, & Jarvis, E.D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, 4, 12.

Rogers, T., Rakinson, D., & McClelland, J. (2004). U-shaped curves in development: A PDP approach. *Journal of Cognition and Development*, 5,137–145.

Thorpe, W.H., & Pilcher, P.M. (1958). The nature and characteristics of subsong. *British Birds*, 51, 509–514.

Todt, D. (1975). Social learning of vocal patterns and modes of their applications in Grey Parrots. *Zeitschrift für Tierpsychologie*, 39, 178–188.

Triana, E., & Pasnak, R. (1981). Object permanence in cats and dogs. *Animal Learning & Behavior*, 9, 135–139

Vehrencamp, S. L., Ritter, A. F., Keever, M., & Bradbury, J. W. (2003). Responses to playback of local vs. distant contact calls in the orange-fronted conure, *Aratinga canicularis*. *Ethology*, 109, 37–54.

Vihman, M.M. (1992). Early syllables and the construction of phonology. In C.A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds), *Phonological development: models, research, implications* (pp. 393–422). Timonium, MD: York Press.

Vihman, M.M. (2014). *Phonological development: The first two years*. (2nd ed.) Malden, MA: Wiley-Blackwell.

Vihman, M. M., & Kunnari, S. (2006). The sources of phonological knowledge. *Recherches Linguistiques de Vincennes*, 35, 133–164.

Warren, D.K., Patterson, D.K., & Pepperberg I.M. (1996). Mechanisms of American English Vowel Production in a Grey Parrot (*Psittacus erithacus*). *The Auk*, 113, 41–58.

Wauquier, S., & Yamaguchi, N. (2013). Templates in French. In Vihman, M. & Keren-Portnoy, T. (Eds), *Child phonology: Whole word approaches, cross-linguistic evidence*, Cambridge : Cambridge University Press.

Weir, R. H. (1962). *Language in the crib*. The Hague: Mouton.