

The Humane Society Institute for Science and Policy

## WBI Studies Repository

2007

# Grey Parrots Do Not Always ‘Parrot’: The Roles of Imitation and Phonological Awareness in the Creation of New Labels from Existing Vocalizations

Irene M. Pepperberg  
*Brandeis University*

Follow this and additional works at: [https://www.wellbeingintlstudiesrepository.org/acwp\\_asie](https://www.wellbeingintlstudiesrepository.org/acwp_asie)



Part of the [Animal Studies Commons](#), [Comparative Psychology Commons](#), and the [Other Animal Sciences Commons](#)

---

### Recommended Citation

Pepperberg, I. M. (2007). Grey parrots do not always ‘parrot’: the roles of imitation and phonological awareness in the creation of new labels from existing vocalizations. *Language sciences*, 29(1), 1-13.

This material is brought to you for free and open access by WellBeing International. It has been accepted for inclusion by an authorized administrator of the WBI Studies Repository. For more information, please contact [wbisr-info@wellbeingintl.org](mailto:wbisr-info@wellbeingintl.org).



# Grey Parrots Do Not Always ‘Parrot’: The Roles of Imitation and Phonological Awareness in the Creation of New Labels from Existing Vocalizations

Irene M. Pepperberg  
Brandeis University  
Harvard University

## KEYWORDS

artificial intelligence, evolution of language, language acquisition, parrot cognition, phonological awareness

## ABSTRACT

*Evidence exists for a form of imitation, vocal segmentation, by a Grey parrot. Data show that the bird understands that his labels are comprised of individual units that can be recombined in novel ways to create a novel referential vocalization; that is, a novel act. Previous data suggested, but could not substantiate, this behaviour. Such evidence implies that a parrot not only has phonological awareness but also demonstrates true imitation rather than mimicry, and has implications for the studies of both the evolution of communicative competence and the development of robotic speech.*

## 1. Introduction

Fairly recently, researchers have begun to view imitation as a complex, advanced behaviour, deserving of detailed investigation and re-evaluation (e.g., Dautenhahn and Nehaniv, 2002, in press; Hurley and Chater, 2005), and a behaviour quite distinct from mimicry, which is generally defined as mindless repetition (e.g., Tomasello and Carpenter, 2005). Imitation, unlike mimicry, primarily involves the intentional copying of an otherwise improbable, novel act (Thorpe, 1963); thus the intentional, referential reproduction of a novel English vocalization by a Grey parrot (*Psittacus erithacus*) is a likely candidate for imitative behaviour. And, given that imitation can also be seen as the integration of a number of familiar actions in novel ways to produce that novel act (e.g., Piaget, 1962; review in Arbib, 2002), of particular interest is what happens when the targeted novel vocalization can be constructed from related elements already in the parrot’s repertoire. This particular type of combinatory behaviour is actually a form of vocal segmentation. Successful segmentation shows that the bird understands that his *existent* labels are comprised of *individual units* that can be *recombined* in novel ways to create novel vocalizations. Previous data suggested, but could not substantiate, this behaviour; current data do just that. Moreover, such evidence implies that a parrot has phonological awareness (sensu Anthony and Francis, 2005) or at

least control of linguistic processing and analysis of linguistic knowledge. The results have implications for other fields, notably the evolution of language and the development of speech in robots and avatars.

Demonstrating segmentation by a Grey parrot would be an important milestone in comparing animal and human cognitive and communicative abilities. Although evolutionarily distant from humans, Grey parrots have already been shown to use elements of English speech referentially (Pepperberg, 1999),<sup>1</sup> and those areas of the human brain responsible for vocal learning and production appear to have analogues (and possible homologues) in parrots (Jarvis et al., 2005). Despite these findings, parrots are still sometimes regarded as mindless mimics. At least two reasons exist for that belief.

One reason, that avian imitation of English speech does not involve intentional, accurate reproduction of human articulatory acts (as far as is possible with parrot anatomy), has been countered previously (e.g., Patterson and Pepperberg, 1994, 1998; Pepperberg, 2002). In brief, although parrot and human anatomy do differ, at least some parrots use a two-tube system and frequency modulation as do humans, and employ their tongues, glottis, and larynx in some of the same ways used by humans to produce vowels and consonants (Warren et al., 1996).<sup>2</sup> For one parrot in the current study, stops exhibit voiced/voiceless, labial, alveolar, and velar groupings, and vowels can be classified with respect to formant structures fairly similar to those of humans, though most of the parrot's variation occurs in the second formant. This parrot is not unique; similar data have been reported by researchers who are teaching their Grey parrots to produce Italian words and phrases (Masin et al., 2005).

Another reason involves segmentation: Only limited evidence exists that parrots, or any animal taught a human communication code, can indeed segment the human code, that is, recombine existing labels intentionally either to describe novel situations or, for example, to produce a phrase to request novel items—rather than simply produce several referential labels that, by chance, appropriately apply to the situation (Fouts and Rigby, 1977; Pepperberg, 1999; Savage-Rumbaugh et al., 1993). Such intentional creativity is, in contrast, common in the earliest stages of normal human language acquisition (de Boysson-Bardies, 1999; Greenfield, 1991; Marschark et al., 1987; Tomasello, 2003). Another form of segmentation, the intentional recombination of existing morphemes or phonemes (parts of words) or their approximations in order to create or reproduce what is for the subject a novel targeted utterance (Greenfield, 1991; Peperkamp, 2003), has not previously been reported in animals; it is not only considered basic to human language development (Carroll et al., 2003), but also a uniquely human trait.

Such phonological awareness, which requires understanding that words are made up of a finite number of sounds that can be recombined into an almost infinite number of patterns (limited only by the constraints of a given language)—the parsing of a complex entity into pieces that are then integrated into a new schema that represents the imitated act (Arbib, 2002)—has additionally been considered a trait that is acquired over time. Children, for example, apparently shift from recognizing and producing words holistically (a simple form of imitation, Studdert-Kennedy, 2002; Arbib, 2005) to recognizing words as being constructed via a rule-based phonology sometime around three years of age or later (Carroll et al., 2003; Vihman, 1996); furthermore, manipulation of individual parts of words is presumed to require development of an internal representation of phonological structure (Byrne and Liberman, 1999). That is, in order to sound out—i.e., to imitate, rather than mimic—a novel label, the child must segment the stream of sound into discrete elements, recognize a match between those elements and elements (or close approximations) that exist in its own repertoire, and then recombine these elements in an appropriate sequence (see Gathercole and Baddeley, 1990; Treiman, 1995; Arbib, 2005). Moreover, children's ability to focus on the sounds of words and sound elements of words rather than solely on word meaning appears to be assisted by training in sound-letter associations (Carroll et al., 2003; Mann and Foy, 2003). Most animals, lacking speech, are never exposed to, nor trained nor tested on, such issues of

phonological awareness or imitation, nor are they expected to have internal representations of phonemes.<sup>3</sup>

Evidence now exists for this form of imitation (vocal segmentation, possibly phonological awareness, but at least metalinguistic awareness) by a Grey parrot: Here I show that my oldest speech-trained subject, Alex, understands that his labels are made of individual phonological units that can be recombined in novel ways to create novel vocalizations. Such evidence implies that parrots not only use English labels referentially, but also understand how such labels are created from independent sound patterns. I also suggest that the bird's ability is indeed a learned behaviour, is not uniquely human, and is dependent upon having considerable experience with both English speech and sound-letter training. My younger birds, lacking such training, do not engage in such behaviour. Of specific interest is that this behaviour occurred in contrast to my parrots' customary patterns of label acquisition, and demonstrates the steps the bird goes through in producing the imitated label.

## **2. Method**

### *2.1. Subjects*

Subjects were two Grey parrots (*P. erithacus*). Alex, 27 years old at the time of this study, had had 26 years of intense training in interspecies communication (Pepperberg, 1999); Arthur, then 3 ½ years old, had had the equivalent of about a year of comparable interspecies communication training. The birds live in a laboratory setting at all times. Their housing and day-to-day care are described in Pepperberg and Wilkes (2004). Using the training technique described below, Alex had previously learned to identify, request, refuse, categorize, and quantify a large number (>100) of objects referentially using English speech sounds (Pepperberg, 1999), and Arthur had already acquired four referential labels (Pepperberg and Wilkes, 2004).

### *2.2. General procedures*

Grey parrots in my laboratory generally learn referential English speech (e.g., to comprehend and produce labels for objects, colors, shapes; to answer questions about concepts of number, category, relative size, absence, same/different) via training with the Model/Rival (M/R) procedure (Pepperberg, 1999). This technique, introduced by Todt (1975) and adapted by Pepperberg (1981), involves three-way interactions between two human speakers and the avian student. While the bird watches, two humans handle an object in which the bird has demonstrated interest; one (the trainer) then questions the other (the parrot's model and rival for the trainer's attention) by using phrases like "What's here?", "What toy?", "What do you want?", etc. The trainer rewards correct responses with the object, thus demonstrating a label's referentiality (the connection between the label and the object to which it refers) and functionality (e.g., showing how the label can be used, as in a request). Humans model errors (e.g., poor pronunciation or other errors similar to the bird's at the time) and demonstrate the consequences of erring by having the trainer say "No!", look away, and briefly remove the object from sight; a correction procedure then follows. The parrot is also questioned and rewarded for an attempt or scolded for an error. Humans exchange roles of trainer and model on a regular basis, thus showing that one individual is not always the questioner and the other the respondent; the bird thus learns to interact with all the humans. Without such role reversal, a bird interacts with only the human acting as trainer (Todt, 1975). By involving the parrot in these interactions, the humans can adjust training to the bird's level (e.g., successively requiring better pronunciation).

This technique was used to train both Arthur and Alex on the label "spool". Arthur was trained first. After Arthur's training, Alex began to show interest in the object, which he had previously ignored. Subsequent

to Arthur's training, Alex, when given the item, began to chew it apart or roll it around his play stand. We therefore decided to initiate M/R training on the object for Alex.

### 2.3. Phoneme training

For several years, Alex had received M/R training to associate the wooden or plastic graphemes B, CH, I, K, N, OR, S, SH, T with their corresponding appropriate phonological sounds (e.g., /bi/ for BI), with the plastic or wooden graphemes as his reward. Although his accuracy was greater than chance (generally about 50%,  $p < .01$ , chance of 1/9), it was never high enough (i.e., ~80%) to claim he had mastered the task.

### 2.4. Taping and sonagraphic analysis

Birds were taped with an AKG-70 microphone directly into an IBM T20 laptop computer; wav. files were edited with Audacity and made into sonagrams by Dr. Diana Reiss (Wildlife Conservation Society, NYC) and Dr. Donald Kroodsma (UMass-Amherst) using *Raven* (Cornell Laboratory of Ornithology). An additional sonagram of Alex's and my formants was made by Dr. Ofer Tchernichovski (CCNY) with *Sound Analysis Pro* (<http://ofer.sci.cuny.cuny.edu>).

## 3. Results

The present data for my oldest subject, Alex, were obtained after our youngest subject, Arthur, had acquired the label spool to refer to wooden or plastic bobbins. The birds' labels usually appear in sessions initially as rudimentary patterns—first a vocal contour, then with vowels, finally with consonants (Patterson and Pepperberg, 1994, 1998) and Arthur's production followed the customary acquisition pattern, that is, beginning with /u:/ ("ooo") and ending with a distinct, fully-formed "spool" (/spu:l/; see Fig. 1A; the IPA transcription is approximate as Arthur used more of a whistle than an actual /p/ and vowel sound). Although Arthur occasionally misidentified the object as "wool" or "wood" (some of his other labels at the time, Pepperberg and Wilkes, 2004), with "wood" sometimes being a correct response for the object's material and "wool" being a reasonable phonological error, he did not consistently use such labels during training and was correct 87.5% on testing (Pepperberg and Wilkes, 2004).

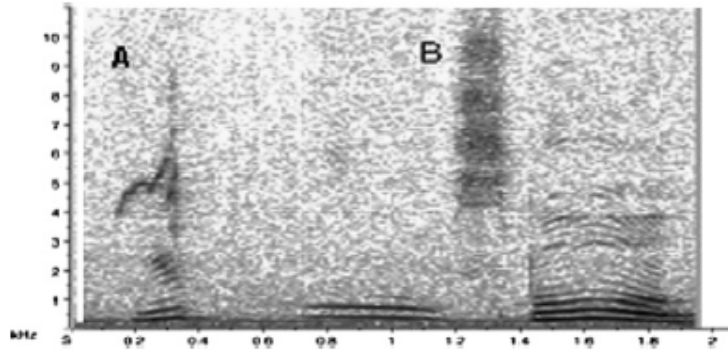
Unlike Arthur, and unlike his usual form of acquisition, Alex, during training after watching Arthur playing with the object, began using a combination of existing phonemes and labels to identify the object: /s/ (trained independently in conjunction with the physical letter, S) and wool, to form "s" (pause) "wool" ("s-wool"; /s-pause-wUl/; Fig. 2).

Note that no labels existed in Alex's repertoire that contained /sp/, nor did he have the labels "pool" or "pull", or any other label that included /Ul/; he did have labels such as "paper", "peach", "parrot", "pick", etc (Pepperberg, 1999). Analysis of Alex's utterances suggest that he produces a viable /p/ via a form of esophageal speech (Patterson and Pepperberg, 1998). He retained this "s-wool" formulation for almost a year of M/R training, with no change whatsoever in the form of his production, although normally only about 25 M/R sessions (at most, several weeks of training) are sufficient for learning a new label (Pepperberg, 1999). The third parrot in the lab, Griffin, who was just beginning training on phonemes, and who heard exactly the same information from Arthur's sessions, did not exhibit this behaviour.

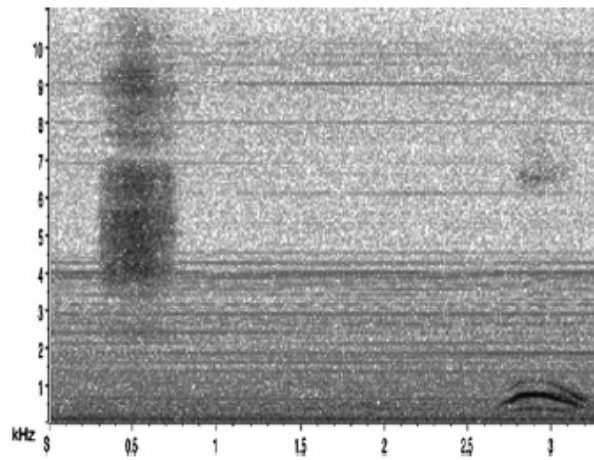
At the end of this year-long period, Alex spontaneously produced "spool", perfectly formed (/spu:l/; see Fig. 3), when I rewarded Arthur with the spool for producing the label. Thus, Alex added the sound—which we hear, sonagraphically view, and transcribe, as—/p/ and also shifted the vowel toward the appropriate /u:/ sound; note how Alex's vowel changes from Fig. 2 to Fig. 3. (Interestingly, both Alex's and my /u:/s are diphthongs,<sup>4</sup> differing slightly from standard American English productions; Patterson

and Pepperberg, 1994.) Because the label “spool” appeared without any intermediary form from that of “s-wool”, no statistical or other analysis of the process of change was possible.<sup>5</sup>

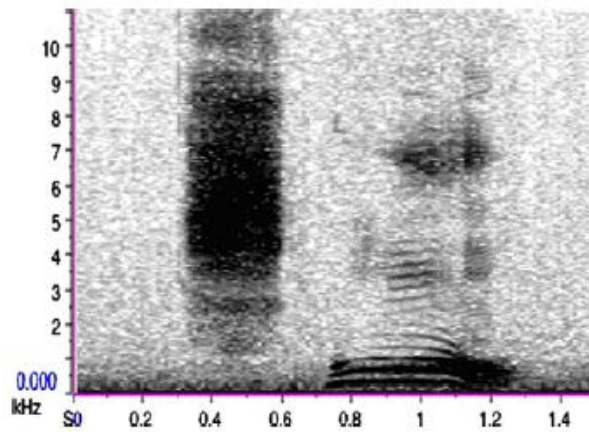
**Fig. 1. Sonagram of (A) Arthur’s “spool”, (B) Pepperberg’s “spool”.**



**Fig. 2. Sonagram of Alex’s “s-wool”.**



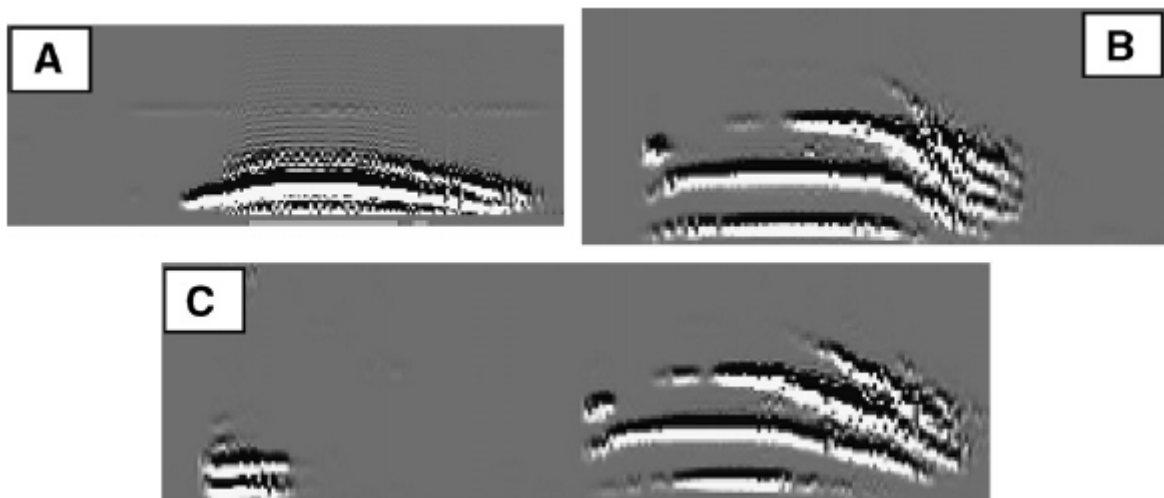
**Fig. 3. Sonagram of Alex’s “spool”.**



**Fig. 3. Sonagram of Alex’s “spool”.**

Alex's and Arthur's productions differ significantly in auditory and sonographic patterns (see Figs. 1A and 3; .wav files on request), such that Alex did not simply learn to mimic Arthur's production; Arthur's utterance incorporated a avian whistle-like quality whereas Alex's utterance sounded distinctly human. Alex's vocal pattern more closely resembles mine (see Fig. 1B), even though I did less than one-tenth of the M/R training. I had, however, done the majority of training on *wool* almost 20 years earlier (Pepperberg, 1999). In general, Alex's formant structures closely approach, although are not identical to, my own (Fig. 4; see Patterson and Pepperberg (1994, 1998) for detailed analyses of the similarities and differences between Alex's and my speech acts; Alex's patterns all closely approach mine, although identity is impossible by virtue of the difference in vocal tract sizes and Alex's lack of lips).<sup>6</sup>

**Fig. 4. A,B: Closeup of (A) Alex's formants for vowel part of "s-wool", (B) Alex's formants for the vowel part of "spool", (C) Pepperberg's formants for the /p/ and vowel part of "spool".**



#### 4. Discussion

As noted above, parrots usually acquire labels by building the sound patterns gradually, beginning with vowels (Patterson and Pepperberg, 1994, 1998). Such behavior may simply reflect the relative ease with which sounds that are more sonorous can be produced relative to those that require, for example, plosive qualities in a subject lacking lips. Nevertheless, completely-formed new labels did occasionally materialize after minimal training and without overt practice (Pepperberg, 1983). In these cases, however (e.g., Alex's production of "carrot" the day after asking us what we were eating, or of the novel label "banerry" to refer to an apple), such utterances appeared fully-formed, with immediacy and no overt practice (Pepperberg, 1999). Even though the label generally contained sounds already in the repertoire (e.g., for "carrot", the /k/ from *key*, the remainder from *parrot*; "banerry" was derived from "banana"—"cherry"), neither my students nor I could convincingly argue that Alex had deliberately parsed labels in his repertoire to match a targeted utterance or to form novel vocalizations. A related argument could be made for Alex's abilities to referentially produce words that form minimal pairs (Patterson and Pepperberg, 1998). So, although Alex could state "Want corn" versus "Want cork", or "Want tea" versus "Want pea" (and refuse the alternatives), which suggests an ability to segment phonemes from the speech stream (somewhat like nonhuman primates; see Newport et al. (2004)), we could not claim that he deliberately parsed these labels when learning to produce them.

That is, we could not claim that he acoustically represented labels as do humans with respect to phonological categories and understood that his labels are made of individual elements that can be

recombined in various ways to produce new ones. Possibly production of “carrot” was potentiated by his already being able to manipulate his vocal tract to produce such sounds, or the new labels were simply created from phonotactically probable sequences involving beginnings and ends of existent labels (Storkel, 2001), or, in the case of “banerry”, from semantic relations. Moreover, data (Patterson and Pepperberg, 1994, 1998) demonstrating that he (a) recognizes small phonetic differences (“tea” vs “pea”) as meaningful, (b) produces initial phonemes differently depending upon subsequent ones (/k/ in “key” vs “cork”), and (c) consistently recombines parts of labels according to their order in existent labels (i.e., combines beginnings of one label with the ends of others—after analyzing over 22,000 vocalizations, we never observed backwards combinations such as “percup” instead of “cupper/copper”; Pepperberg et al., 1991), merely imply but did not prove that he engages in such *top-down* processing (Ladefoged, 1982).

Even the two closest behaviour patterns previously reported that suggest some form of label parsing (Pepperberg, 1990; Pepperberg et al., 1991), which both involve solitary sound play, differ from the kind of bootstrapping described here. In one behaviour, Alex produced strings such as *mail chail benail* in private practice before producing the targeted, trained label *nail* (Pepperberg et al., 1991). The *nail* situation differs from bootstrapping in that the combinations of phonemes did not seem to be a deliberate attempt to create a new label from specific sound patterns that resembled the target, but rather to be deliberate play within a range of existent patterns in an attempt to hit on a correct pairing that matched some remembered template. That is, Alex’s behaviour demonstrated an understanding of the combinatory nature of his utterances, but did not show that he understood how to segment the novel targeted vocalization exactly, then match its components to those in his repertoire in order to create trained label. In the second behaviour (Pepperberg, 1990), Alex produced strings such as *grape, grain, chain, cane* in the absence of specific objects but in the presence of his trainers. Although these labels could quickly be referentially mapped onto physical objects, we had no reason to believe that production of such strings was intentional, other than to gain the attention of trainers. Note that with the exception of *grape*, such labels would not have been used in the laboratory. The rhyme awareness demonstrated in these behaviour patterns, although separate from phoneme awareness, is still considered closely aligned to children’s language skills (see Mann and Foy, 2003), and again supports the argument that Alex views his labels as being constructed from individual sound patterns.

The current data, when taken in combination with previous evidence, however, suggests that at least one parrot, much like a child, can actually apply a phonological rule derived from knowledge of its repertoire: recognize that sounds such as “car” and “pet” can be recombined for use in identifying and creating the basis for a totally distinct object—carpet—whose label has no referential correlation to the original utterances.<sup>7</sup> That is, Alex appeared to form the closest match based on segmentation and onset + nucleus + coda (Storkel, 2002). Arguably, the data presented here could be considered stronger if Alex had had the labels *pull* or *pool* in his repertoire and had initially produced either “s (pause) pull” or “s (pause) pool”. Note, however, that /p/ is particularly difficult for a parrot, lacking lips, to produce (Patterson and Pepperberg, 1998); Alex’s first attempts at “peach”, for example, sounded like “cheech” (Neal, 1996), and, as noted above, Patterson and Pepperberg (1998) have suggested that he may be using a form of esophageal speech for /p/. I thus believe that his production of “s-wool” is actually more important, because, not having exact matches, he took the closest, readily-available sounds in his repertoire (i.e., “wool” is the only label out of approximately 90 documented in his repertoire that resembles “spool”) to form the initial attempts at a novel vocalization, and by so doing, made the process transparent to his human trainers.

Another interesting issue exists concerning Alex’s behaviour, and addresses the issue of whether he has simply shown a sensitivity to sound similarity. Exactly because of the difficulty of producing /p/, Alex may



have used “s-pause-wool” as a way of initiating the vocalization such that two known utterances provided the overall structure and the pause was a place filler, somewhat like that occasionally used by young children, until he could learn how to insert the /p/ and adapt the vowel. Specifically, Peters (2001) suggests that children use certain sounds as fillers (a “holding tank”) to preserve the number of syllables or the prosodic rhythm of the target vocalization until the standard form is learned (note also Leonard, 2001). Even though the work of Peters and colleagues primarily refer to the earliest forms of grammar rather than labels (but note examples in Lleo´, 2001), and that Alex used a pause, rather than another phoneme, his behaviour suggests (but, of course, does not prove), that he had an awareness of the need for something additional and somewhat different to complete the vocalization. Simply omitting or closing the gap—and responding on the basis of sound similarity—would have produced /swU/ (“swull”), not /swu:l/ (“swoooo”).

One might, of course, question whether this single instance of combinatory behavior qualifies as evidence of phonological awareness in a nonhuman. Arguably, Alex may have applied a phonological rule for combining utterances without truly understanding the basis for the rule; that is, he may have engaged in some metalinguistic analysis (e.g., about beginnings and ends of labels), but without awareness that he was working with phonemes. Such an argument could, indeed, be made for the labels that he produced in the absence of referents (i.e., in apparent sound play, Pepperberg, 1990), but the specificity and consistent use of the “s-wool” combination argues against such an alternative explanation, as well as against that of “babble-luck” (a fortuitously correct but accidental combination, Thorndike, 1943). Here, Alex had to have discriminated and extracted the appropriate speech sounds of the target label “spool”, generalized these to the closest related items in his repertoire, fit the existent sounds together—apparently including a pause to maintain spacing for an absent sound—in a particular serial order so as to add to his lexicon, and additionally had to link the novel phonology referentially with a specific item. The uniqueness of the behaviour (i.e., the single reported label) is also not remarkable: For the past several years, Alex’s training has concentrated on concept, rather than label, acquisition, and he has had few opportunities to engage in novel label learning. Interestingly, in recent training on the label “seven”, his first attempts have been “ss. . .nnnn”, which also suggest some level of phonological awareness.

It is now possible to return to the initial argument, that Alex’s vocal segmentation provides evidence for true imitation, rather than mimicry. ‘Mere’ mimicry can be defined as the purposeless duplication of an act (for a bird, it would be rote reproduction of human speech without referential content), behaviour that lacks cognitive complexity and intentionality (e.g., Tomasello and Carpenter, 2005). But if an act is performed because the imitator understands its purpose—to reach a goal, be it an object or intentional communication, otherwise impossible to obtain—then the act is intentional, complex, likely indicates cognitive processing, and provides evidence for true imitation. As stated in Section 1, Alex’s data demonstrate that he understands that his existent labels are comprised of *individual units* that can *intentionally be recombined* in novel ways to create referential, novel vocalizations.

Whether such data can be used to argue for a parrot’s understanding of a phonetic ‘grammar’ (e.g., Fitch and Hauser, 2004) is unclear: Although the data suggest that Alex can generate novel meaningful labels from a finite set of elements, the rule system he demonstrated was relatively limited. Nevertheless, the data presented here add another intriguing parallel between Alex’s and young children’s early label acquisition (Pepperberg, 1999). For children, manipulation of individual parts of a word implies the existence of internal representations of words as divisible units, and normal children proceed in a fairly standard manner from babbling to full language. Alex does not have and will not likely reach the level of any young child—that is, in terms of grammar go beyond use of simple sentence frames such as “I want X” and “I wanna go Y”, where X and Y are appropriate object or location labels—but any strides that a bird makes toward language-like ability—such as, for example, comprehending recursive conjunctive

sentences<sup>8</sup> or demonstrating the kind of vocal segmentation described here—helps us understand the similarities and differences between humans and nonhumans. Particularly, how do these data affect the idea that the evolution of human language must be based on phylogeny—that anything language-like in nonhumans must be a function of their common ancestry with humans and that only nonhuman primates would be expected to exhibit precursors to human language (e.g., Savage-Rumbaugh et al., 1993)?

Although Alex's abilities are clearly not isomorphic with human language, our data (as well as that of previous studies; see Pepperberg and Shive, 2001) demonstrate that elements of linguistic behaviour are not limited to primates, nor are the neurological systems underlying such behaviour. Avian neuroanatomy and its relation to the mammalian line is, however, not yet well enough understood to determine specific parallels among oscine, psittacine, and mammalian structures, although significant progress is being made (e.g., Jarvis et al., 2005). Nevertheless, given the evolutionary distance between parrots and primates, the search for and arguments concerning responsible neural substrates and common behaviour should be approached with care and not be restricted to the primate line. The data presented here, plus our knowledge of avian vocal learning, of how social interaction affects such learning, and of birds' advanced cognition (Clayton et al., 2005; Kenward et al., 2005; Kroodsma and Miller, 1996; Pepperberg, 1999; Weir et al., 2002), all suggest *Aves* as an important model for determining the evolutionary pressures responsible for—and in developing testable theories about—complex communication systems, particularly those involving vocal learning.

Finally, these findings may also be of use in two ways for computer scientists who are trying to develop speech skills in their avatars and robots. First, as Patterson and Pepperberg (1994, 1998) have demonstrated, Alex produces most of his utterances with little variation in his first formant, and most of the variation in his second (and possibly third). Thus speech modelling may be simplified if based on avian, rather than human, productions (note Schwarz et al., 2001). Second, Alex's pattern of acquisition might suggest how approximations and iterations can be used for the construction of novel speech sounds from existent programmed vocalizations, providing additional assistance to existent algorithms (see Higashimoto and Sawada, 2002; Nishikawa et al., 2002; Yoshikawa et al., 2003a,b).<sup>9</sup>

In sum, I suggest that Alex's training on both referential labeling and sound-letter association has engendered levels of phonological awareness, vocal segmentation, and imitation that need to be addressed when arguing for (a) human uniqueness and (b) the exclusive use of humans as the bases for computational models.

## Notes

<sup>1</sup> No claim is made that Alex's speech is isomorphic with human language (e.g., Alex cannot discourse about the weather), only that the elements that he does produce have been documented as being used referentially; labels are both understood and used in contexts that differ from and extend beyond training conditions.

<sup>2</sup> Note that such data actually support a motor theory of speech perception; a detailed discussion of this issue, along with the relevance of motor neurons in birds for vocal learning is forthcoming (Pepperberg, in press, in preparation); for now, see Pepperberg (2005).

<sup>3</sup> Nonhuman primates have been trained and tested on their ability to segment human speech sounds (e.g., Newport et al., 2004), but not on sound-letter associations or on productive recombination of speech elements.

<sup>4</sup> Alex's and my /u:/ both consists of two units; see Patterson and Pepperberg (1994) for details.

<sup>5</sup> For Wav. forms, contact [impepper@media.mit.edu](mailto:impepper@media.mit.edu). Production of "spool" first occurred during a taping session with the BBC, and after production of "s-wool" during the same session; thus the vocalization was evaluated by observers who had had no possible previous connection with training sessions.

<sup>6</sup> See Beckers et al. (2004) for discussion of tongue placement and formation of true formants in parrot vocalizations.

<sup>7</sup> Clearly, carpet is not pronounced as car + pet, but car and pet can be combined and then modified to form the targeted utterance.

<sup>8</sup> For example, given various trays each holding seven objects of several colors, shapes, and materials, Alex can respond to queries of "What object/material is color-A and shape-B?" versus "What shape is color-A and object/material-C?" versus "What color is shape-B and object/material-C?" (Pepperberg, 1992).

<sup>9</sup> Note that 'parrot-like teaching' for these authors is entirely different from our M/R technique, and does not at all involve parrots.

## Acknowledgements

I thank Drs. Diana Reiss (Wildlife Conservation Society, NYC) and Donald Kroodsma (UMass-Amherst) for creating the sonagrams on Raven, Dr. Ofer Tchernichovski (CCNY) for the sonagram using *Sound Analysis Pro*, and Dr. Diane Patterson for alerting me to the issue of 'filler' phonemes. The research was supported by donors to *The Alex Foundation*, with special thanks to the Pearl Family Foundation, Dr. Janice Boyd, Kathryn McAdams, Greg Harrison and family (Harrison's Bird Food), the Makioka Foundation, Janet and Kip Trumbule, Greg Lamorte, Nancy Clark and Bill Broach, LeighAnn and Carl Hartsfield, Nancy Chambers, Michael and Alex Shuman, Bob Goodale and Deb Rivel Goodale.

## References

- Anthony, J.L., Francis, D.J., 2005. Development of phonological awareness. *Current Directions in Psychological Science* 14, 255–259.
- Arbib, M.A., 2002. The mirror system, imitation, and the evolution of language. In: Dautenhahn, K., Nehaniv, C.L. (Eds.), *Imitation in Animals and Artifacts*. MIT Press, Cambridge, MA, pp. 229–279.
- Arbib, M.A., 2005. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavior & Brain Sciences* 28, 105–167.
- 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.
- Byrne, B., Liberman, A.M., 1999. Meaninglessness, productivity and reading: Some observations about the relation between the alphabet and speech. In: Oakhill, J., Beard, R. (Eds.), *Reading Development and the Teaching of Reading*. Blackwell, Oxford, pp. 157–174.
- Carroll, J.M., Snowling, M.J., Hulme, C., Stevenson, J., 2003. The development of phonological awareness in preschool children. *Developmental Psychology* 39, 913–923.

Clayton, N.S., Dally, J., Gilbert, J., Dickinson, A., 2005. Food caching by western scrub-jays (*Aphelocoma californica*) is sensitive to the conditions at recovery. *Journal of Experimental Psychology: Animal Behavior Processes* 31, 115–124.

de Boysson-Bardies, B., 1999. *How Language Comes to Children*. MIT Press, Cambridge, MA.

Dautenhahn, K., Nehaniv, C. (Eds.), 2002. *Imitation in Animals and Artifacts*. MIT Press, Cambridge, MA.

Dautenhahn, K., Nehaniv, C., (Eds.), in press. *Imitation and Social Learning in Robots, Humans and Animals: Behavioural, Social and Communicative Dimensions*, Cambridge University Press, Cambridge, UK.

Fitch, W.T., Hauser, M.D., 2004. Computational constraints on syntactic processing in a nonhuman primate. *Science* 303, 377–380.

Fouts, R., Rigby, R., 1977. Man–chimpanzee communication. In: Sebeok, T. (Ed.), *How Animals Communicate*. Indiana University Press, Bloomington, IN, pp. 1034–1054.

Gathercole, S.E., Baddeley, A.D., 1990. The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology* 81, 439–454.

Greenfield, P., 1991. Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavior & Brain Sciences* 14, 531–595.

Higashimoto, T., Sawada, H., 2002. Speech production by a mechanical model: construction of a vocal tract and its control by neural network. In: *Proceedings of the 2002 IEEE International Conference on Robotics and Automation*, pp. 3858–3863.

Hurley, S., Chater, N. (Eds.), 2005. *Perspectives on Imitation: from Neuroscience to Social Science*. MIT Press, Cambridge, MA.

Jarvis, E.D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., Medina, L., Paxinos, G., Perkel, T., Shimizu, T., Striedter, G., Wild, J.M., Ball, G.F., Dugas-Ford, J., Durand, S.E., Hough, G.E., Husband, L., Kubikova, L., Lee, D.W., Mello, C.V., Powers, A., Siang, C., Smulders, T.V., Wada, K., White, S.A., Yamamoto, K., Yu, J., Reiner, A., Butler, A.B., 2005. Avian brains and a new understanding of vertebrate evolution. *Nature Reviews Neuroscience* 6, 151–159.

Kenward, B., Weir, A.A.S., Rutz, C., Kacelnik, A., 2005. Tool manufacture by naive juvenile crows. *Nature* 433, 121.

Kroodsma, D.E., Miller, E.H. (Eds.), 1996. *Ecology and Evolution of Acoustic Communication in Birds*. Cornell University Press, Ithaca.

Ladefoged, P., 1982. *A Course in Phonetics*. Harcourt Brace Jovanovitch, San Diego, CA.

Leonard, L.B., 2001. Fillers across languages and language abilities. *Journal of Child Language* 28, 257–261.

Lleó, C., 2001. Early fillers: undoubtedly more than phonological stuffing. *Journal of Child Language* 28, 262–265.

Mann, V.A., Foy, J.G., 2003. Phonological awareness, speech development, and letter knowledge in preschool children. *Annals of Dyslexia* 53, 149–173.

- Marschark, M., Everhart, V.S., Martin, J., West, S.A., 1987. Identifying linguistic creativity in deaf and hearing children. *Metaphor & Symbolic Activity* 2, 281–306.
- Masin, S., Lenti Boero, D., Massa, R., Bottoni, L., 2005. Acoustic characteristics of vowel-like sounds of an African Grey parrot (*Psittacus erithacus*). Poster Presented at the Avian Brain Conference, XXIXth Int'l. Ethological Conference, August.
- Neal, K.B., 1996. The development of a vocalization in an African Grey parrot (*Psittacus erithacus*). Unpublished senior thesis, University of Arizona, Tucson, AZ.
- Newport, E.L., Hauser, M.D., Spaepen, G., Aslin, R.N., 2004. Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology* 49, 85–117.
- Nishikawa, K., Imai, A., Ogawara, T., Takanobu, H., Mochida, T., Takanishi, A., 2002. Speech planning of an anthropomorphic talking robot for consonant sounds production. In: Proceedings of the 2002 IEEE International Conference on Robotics and Automation, pp. 1830–1835.
- 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. Acoustic and articulatory correlates of stop consonants in a parrot and a human subject. *Journal of the Acoustical Society of America* 103, 2197–2215.
- Peperkamp, S., 2003. Phonological acquisition: Recent attainments and new challenges. *Language & Speech* 46, 87–113.
- Pepperberg, I.M., 1981. Functional vocalizations by an African Grey Parrot (*Psittacus erithacus*). *Zeitschrift für Tierpsychologie* 55, 139–160.
- Pepperberg, I.M., 1983. Cognition in the African Grey parrot: Preliminary evidence for auditory/vocal comprehension of the class concept. *Animal Learning & Behavior* 11, 179–185.
- Pepperberg, I.M., 1990. Referential mapping: A technique for attaching functional significance to the innovative utterances of an African Grey parrot (*Psittacus erithacus*). *Applied Psycholinguistics* 11, 23–44.
- Pepperberg, I.M., 1992. Proficient performance of a conjunctive, recursive task by an African Grey parrot (*Psittacus erithacus*). *Journal of Comparative Psychology* 106, 295–305.
- Pepperberg, I.M., 1999. *The Alex Studies*. Harvard University Press, Cambridge, MA.
- Pepperberg, I.M., 2002. Allospecific referential speech acquisition in Grey parrots (*Psittacus erithacus*): evidence for multiple levels of avian vocal imitation. In: Dautenhahn, K., Nehaniv, C.L. (Eds.), *Imitation in Animals and Artifacts*. MIT Press, Cambridge, MA, pp. 109–131.
- Pepperberg, I.M., 2005. Insights into vocal imitation in African Grey parrots (*Psittacus erithacus*). In: Hurley, S., Chater, N. (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science*. MIT Press, Cambridge, MA, pp. 242–262.
- Pepperberg, I.M., in press. Emergence of linguistic communication: Studies on Grey parrots. In: Lyon, C., Nehaniv, C.L., Cangelosi, A. (Eds.), *Emergence and Evolution of Linguistic Communication*. Springer.
- Pepperberg, I.M., in preparation. *Learning to Communicate*. Harvard University Press, Cambridge, MA.

- 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., Shive, H., 2001. Simultaneous development of vocal and physical object combinations by a Grey Parrot (*Psittacus erithacus*): Bottle caps, lids, and labels. *Journal of Comparative Psychology* 115, 376–384.
- Pepperberg, I.M., Wilkes, S.R., 2004. Lack of referential vocal learning from LCD video by Grey Parrots (*Psittacus erithacus*). *Interaction Studies* 5, 75–97.
- Peters, A.N., 2001. Filler syllables: What is their status in emerging grammar? *Journal of Child Language* 28, 229–242.
- Piaget, J., 1962. *Play, Dreams and Imitation in Childhood*. Norton, New York.
- Savage-Rumbaugh, S., Murphy, J., Sevcik, R.A., Brakke, K.E., Williams, S.L., Rumbaugh, D.M., 1993. Language comprehension in ape and child. *Monographs of the Society for Research in Child Development* 233, 1–258.
- Schwarz, J.-L., Boë, L.-J., Bessière, P., 2001. Speech robotics, a computational framework for the study and modelling of speech acquisition and evolution. Available from: <[www.icp.inpg.fr/OHLL/lesPubliRapports/article\\_Steeles\\_2001.pdf](http://www.icp.inpg.fr/OHLL/lesPubliRapports/article_Steeles_2001.pdf)>.
- Storkel, H., 2001. Learning new words II: Phonotactic probability in verb learning. *Journal of Speech, Language, & Hearing Research* 44, 1312–1323.
- Storkel, H., 2002. Restructuring of similarity neighbourhoods in the developing mental lexicon. *Journal of Child Language* 29, 251–274.
- Studdert-Kennedy, M., 2002. Mirror neurons, vocal imitation, and the evolution of particulate speech. In: Stamenov, M.I., Gallese, V. (Eds.), *Mirror Neurons and The Evolution of Brain and Language*. John Benjamins, Amsterdam, Netherlands, pp. 207–227.
- Thorndike, E.L., 1943. *Man and His Works*. Harvard University Press, Cambridge, MA.
- Thorpe, W.H., 1963. *Learning and Instinct in Animals*, 2nd ed. Harvard University Press, Cambridge, MA.
- Todt, D., 1975. Social learning of vocal patterns and modes of their applications in Grey parrots. *Zeitschrift für Tierpsychologie* 39, 178–188.
- Tomasello, M., 2003. *Constructing a Language*. Harvard University Press, Cambridge, MA.
- Tomasello, M., Carpenter, M., 2005. Intention reading and imitative learning. In: Hurley, S., Chater, N. (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science*. MIT Press, Cambridge, MA, pp. 135–148.
- Treiman, R., 1995. Errors in short-term memory for speech: A developmental study. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 21, 1197–1208.
- Vihman, M.M., 1996. *Phonological Development: The Origins of Language in the Child*. Blackwell, Malden, MA.

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

Weir, A.A.S., Chappell, J., Kacelnik, A., 2002. Shaping of hooks in New Caledonian crows. *Science* 297, 981.

Yoshikawa, Y., Asada, M., Hosoda, K., Koga, J., 2003a. A constructivist approach to infants' vowel acquisition through mother–infant interaction. *Connection Science* 15, 245–258.

Yoshikawa, Y., Koga, J., Asada, M., Hosoda, K., 2003b. Primary vowel imitation between agents with different articulation parameters by parrot-like teaching. *Proceedings of the 2003 International Conference on Intelligent Robots and Systems*.