LACUS FORUM XXXIV

Speech and Beyond

EKU

GREY PARROT VOCAL LEARNING: CREATION OF NEW LABELS FROM EXISTING VOCALIZATIONS AND ISSUES OF IMITATION

Irene M. Pepperberg
Harvard and Brandeis Universities

Imitation is often considered a prerequisite for communication, and particularly so for interspecies communication, in the sense that the ability to reproduce signals, signs, or symbols in a given context suggests agreement about their reference between the model (in this case, various humans) and the imitator (here a Grey parrot, *Psittacus erithacus*). Nevertheless, considerable confusion exists about the term "imitation," in and of itself, which must be clarified before any discussion of its role in communication. The first step in such clarification is to separate "imitation" from "mimicry," the latter being the mindless, nonreferential repetition often associated with the term "parroting," rather than the intentional, referential use of nonspecies-specific (i.e., heterospecific or allospecific) speech elements by a nonhuman. Imitation has also been defined, notably by Thorpe (1963), as the intentional copying of an otherwise improbable, novel act and, in some cases (e.g., Arbib 2005), as the integration of several familiar actions in novel ways to produce that novel act. Thus, in this paper, I review arguments that the intentional, referential reproduction of novel English vocalizations by a Grey parrot, Alex, likely represents imitative behavior, particularly when the targeted novel vocalizations are constructed from related elements already in his repertoire (i.e., segmentation); I also discuss consequences of this imitative behavior in terms of evaluating this bird's communicative competence (Pepperberg 2007a, b). Previously, Alex had been shown to label over 50 exemplars, 7 colors, 5 shapes, quantities to 6, 3 categories (color, shape, material) and use "no," "come here," "wanna go X" and "want Y" (X and Y are appropriate location or item labels). He combined labels in simple ways to identify, request, comment upon, or refuse more than 100 items and alter his environment. He processed queries to judge category, relative size, quantity, presence or absence of numerical sets and similarity/difference in attributes, and to show label and number comprehension (Pepperberg 1999, 2006). He semantically separated labeling from requesting. Alex had also been trained on phonemes: He associated alphabet letters B, CH, I, K, N, OR, S, SH, T with corresponding appropriate phonological sounds (e.g., /bi/ for BI), receiving the plastic or wooden letters as his reward (Pepperberg 2007a). Alex's abilities, advanced though they were, could not qualify him as having acquired a human language; nevertheless, I will argue that he achieved a distinctive form of interspecies communication, including imitative behavior.

Two arguments have, however, been proposed against interpreting Alex's behavior as true imitation, which also must be addressed before any further discussion (Pepperberg 2007a, b). One argument is that avian imitation of English speech does not involve inten-

Reprinted from LACUS Forum 34: Speech and Beyond, ed. by Patricia Sutcliffe, Lois Stanford, and Arle Lommel. Houston, TX: LACUS. 2008

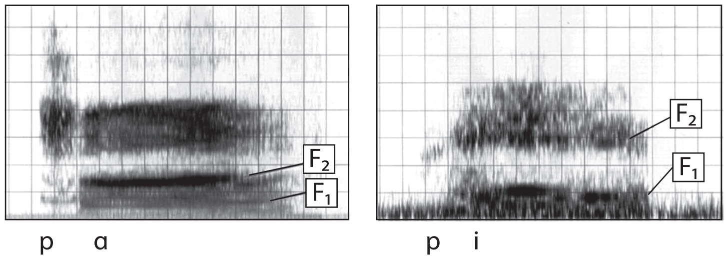

Figure 1. *Alex's production of "pah" /pa/, his label for pasta and for "pea" /pi/, his label for a green pea.*

tional, accurate reproduction of human articulatory acts. The second argument is that non-humans are incapable of segmentation.

The first argument has already been countered (Pepperberg 2007a, b): as described in Patterson and Pepperberg (1994, 1998), Alex's parrot anatomy prevents him from exactly reproducing human articulatory acts, but he (though maybe not all parrots) uses a two-tube system and frequency modulation as do humans, and employs his tongue, glottis, and larynx in some of the same ways used by humans to produce vowels and consonants (Warren, Patterson & Pepperberg 1996). His stops exhibit voiced/voiceless, labial, alveolar, and velar groupings; his vowels can be classified with respect to formant structures similar to those of humans, though most of his variation occurs in the second formant (see **Figure 1**). That is, his speech is not simply the result of, for example, sine wave interference as proposed by Lieberman (1984), but shows formants like those of his trainer.

Countering the second argument, by claiming that Alex is capable of vocal segmentation—a special form of vocal combinatory behavior—would imply that he recognizes that his existent labels are formed of individual morphemes or phonemes that can be combined in novel ways to create what are for the subject novel vocalizations (e.g., Greenfield 1991, Peperkamp 2003), and would also demonstrate phonological awareness (Pepperberg 2007a). Such behavior is not only considered basic to human language development (Carroll *et al.* 2003), but also a uniquely human trait: most animals, lacking speech, are never exposed to, nor trained nor tested on, issues of phonological awareness or imitation, nor are they expected to have internal representations of phonemes that would allow for such combinatorial behavior (Pepperberg 2007a). Even in children, such behavior is not considered innate: Children, for example, apparently shift from recognizing and producing words holistically (a simple form of imitation, see Studdert-Kennedy 2002, Arbib 2005) to recognizing words as being constructed via a rule-based phonology 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 & Liberman 1999). That is, in order to sound out—i.e., to imitate, rather than mimic—a novel label, children must segment the sound stream into discrete elements, recognize a match between those elements and elements (or close approximations) that exist in their own repertoires, and then recombine these elements in an appro-

Reprinted from LACUS Forum 34: Speech and Beyond, ed. by Patricia Sutcliffe, Lois Stanford, and Arle Lommel. Houston, TX: LACUS. 2008

priate sequence (see Gathercole & 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 & Foy 2003). Until now, little evidence exists for any type of segmentation in animals, even a less advanced form involving combination and/or recombination of whole labels to describe novel situations. The few existent incidents—apes' "water bird" for a swan, "cry hurt food" for a radish (Fouts & Rigby 1977), dolphins' "ring-ball" during simultaneous play with two items (Reiss & McCowan 1993)—have been considered as descriptors of the entire situation rather than as specific combinations to denote one element. In this paper, I review evidence for Alex's segmentation (Pepperberg 2007a, b) and present new data confirming this ability.

1. EXPERIMENTAL DESIGN.

1.1. SUBJECTS. The study involved two Grey parrots: Alex, then 27 years old and with 26 years of human interaction and training (see Pepperberg 1999, 2007a); and Arthur, then only 4.5 years old. Although Arthur had had about 3.5 years of intense human interaction, he had the equivalent of only about a full year of communication training (i.e., he knew "tickle," "hello," a generic "want some," and two object labels; Pepperberg & Wilkes 2004). Housing is described in Pepperberg and Wilkes (2004).

1.2. TRAINING. Arthur and later Alex were trained via the standard Model/Rival (M/R) technique (Pepperberg 1981, see also Todt 1975) to produce the label "spool" in response to wooden bobbin. Briefly, this technique uses three-way social interactions among two humans and a parrot to demonstrate a vocal behavior to be learned. The parrot observes two humans interacting as they handle and speak about one or more objects. One trainer presents objects and queries the other human about these items, using expressions such as "What's here?," "What color?," giving praise and transferring the named object to the human partner as a reward for correct answers, thereby providing a one-to-one correlation between object and label. Incorrect responses are punished by scolding and by temporarily removing items from sight. Thus the second human serves both as a model for the parrot's responses and its rival for the trainer's attention, and illustrates the consequences of errors. The model must try again or talk more clearly if the response was deliberately made incorrectly or garbled; that is, the model is subject to the process of corrective feedback, which the bird observes. The parrot is also included in the interactions: it is queried and rewarded for successive approximations to correct responses and training is adjusted to its performance level. Roles of trainer and model are also interchanged, emphasizing that a questioner is sometimes a respondent and demonstrating that the procedure can effect environmental change. Role reversal also counteracts an earlier methodological problem: birds whose trainers always maintained their respective roles responded only to the human questioner (Todt 1975). With this technique, birds will respond to, interact with, and learn from any human.

Reprinted from LACUS Forum 34: Speech and Beyond, ed. by
Patricia Sutcliffe, Lois Stanford, and Arle Lommel. Houston, TX: LACUS. 2008

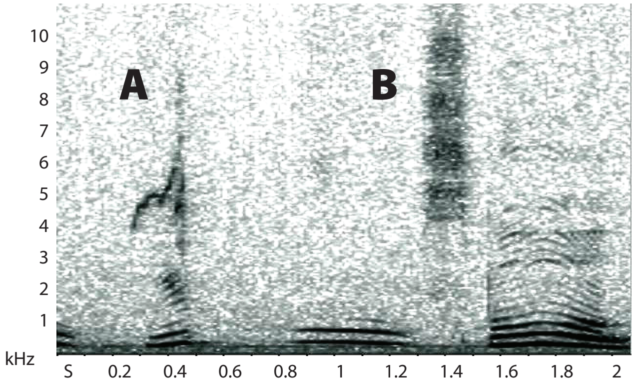

Figure 2. *(a) Arthur's "spool" compared to (b) Pepperberg's "spool" (from Pepperberg 2007a).*

2. RESULTS. Arthur's acquisition initially followed the general pattern for birds in my lab (Patterson & Pepperberg 1994, 1998; Pepperberg 2007a, b). He began with the vowel, /u/ ("ooo"), added the fairly simple consonant /l/, and then, because production of a human /p/ is troublesome without lips, he had difficulty with /p/. Unlike Alex, who learned to produce /p/ apparently via esophageal speech (Patterson & Pepperberg 1998), Arthur's solution was to produce a whistled, not plosive, /p/ in /sp/ (see **Figure 2**; Pepperberg 2007a).

His behavior was similar to what Lieberman (1984:156) had predicted for parrot "speech." Specifically, Lieberman (1984) argued that birds could not reliably produce the same formant structure as humans, but rather, as noted above, produce whistles that, via interference patterns that create energy at defined frequencies, are translated by the human ear into speech-like sounds. Note, however, that only /sp/ followed the whistled pattern whereas the /u/ (which could also easily have been whistled) and /l/ clearly resembled human speech (Pepperberg 2007a), and that previous research (Patterson & Pepperberg 1994, 1998) revealed formant structure for all vowels and stop consonants (/p/, /b/, /d/, /g/, /k/, /t/) for Alex.

After observing the attention that Arthur received for labeling the wooden bobbin, Alex began to show interest in the item, and he received M/R training to produce the label. For "spool," unlike Arthur and unlike his usual form of acquisition, Alex began using a combination of existing phonemes and labels to identify the object: /s/ (unvoiced, trained in conjunction with the alphabet letter, S) and *wool*, to form "s" (pause) "wool" ("s-wool", i.e. /s-pause-wʊl/; see **Figure 3**; Pepperberg 2007a). The pause seemed to provide space for the absent (and difficult) /p/ (possibly as a filler phoneme, preserving the number of syllables or prosodic rhythm of the targeted vocalization; see Leonard 2001, Peters 2001). No prior labels existed in Alex's repertoire containing /sp/, nor did his repertoire include "pull" or "pool," nor any label including /ul/. He did know "paper," "peach," "parrot," "pick," etc. and "shape" and "sich" (six); thus, technically, /p/, /sh/ and /s/ but not /sp/ were available. He knew /u/ from labels such as "two" and "blue" (Pepperberg 2007a). Note that both

Reprinted from LACUS Forum 34: Speech and Beyond, ed. by Patricia Sutcliffe, Lois Stanford, and Arle Lommel. Houston, TX: LACUS. 2008

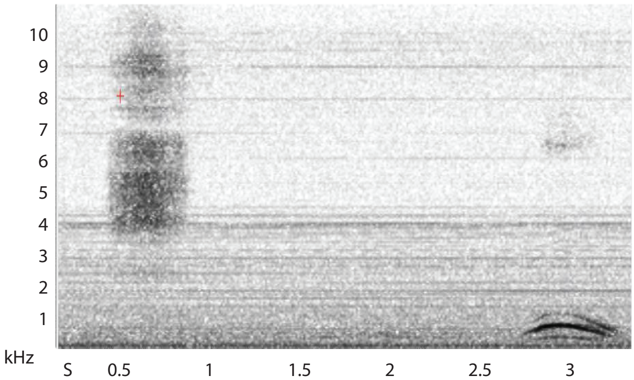

Figure 3. *Alex's "s-wool" (/s-pause-wUl/), (from Pepperberg 2007a).*

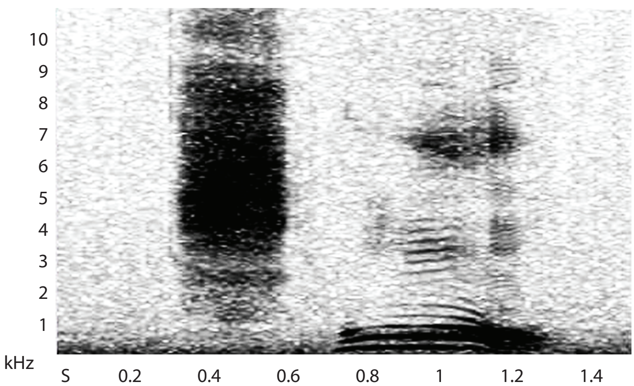

Figure 4. *Alex's "spool" (/spul/) (from Pepperberg 2007a).*

Alex's and Pepperberg's /p/, when analyzed for VOT (voice onset time), fall solidly into the voiceless category and are distinct from the voiced /b/ (Patterson & Pepperberg 1998).

Alex retained "s-wool" for almost a year, even though usually 20–25 modeling sessions (at most several weeks of training) enable learning of a new label with existent phonemes (Pepperberg 1999). At the end of this year-long period, he spontaneously produced a perfectly formed "spool" (/spul/). Thus, he added the /p/ where there had been a clear space and also shifted the vowel toward the appropriate /u/ sound (see **Figure 4** and **Figure 5**, overleaf, Pepperberg 2007a).

A comparison of **Figures 2** and **4** shows that Arthur's and Alex's productions differ significantly in acoustic and sonagraphic patterns. Alex clearly did not imitate or mimic Arthur. Arthur's utterance had a clear avian whistle-like quality; Alex's utterance sounded distinctly human. Alex's utterance clearly resembled that of Pepperberg (**Figure 2B**), even though students had performed 90% of the training. **Figure 5** (overleaf) highlights how

Reprinted from LACUS Forum 34: Speech and Beyond, ed. by Patricia Sutcliffe, Lois Stanford, and Arle Lommel. Houston, TX: LACUS. 2008

Figure 5. *(a) Alex's /U/; (b) Alex's /u/, (c) part of Pepperberg's /spu/ (from Pepperberg 2007a).*

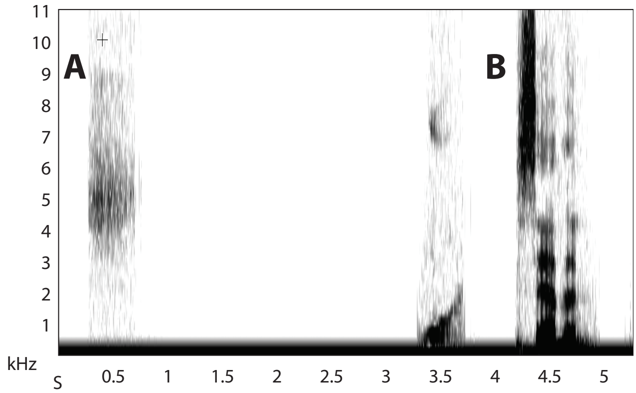

Figure 6. *(a) Alex's "s-one" /s-pause-wən/ followed by (b) Pepperberg's "seven" /sEvIn/*

Alex's vowel section changed from "s-wool" /U/ to "spool" /u/ to resemble that of Pepperberg.

Whether Alex's shift from /U/ to /u/ was gradual or not is unknown. Unlike a previous laboratory situation in which Alex was alone for specific periods each day to enable monitoring of his solitary practice (Pepperberg, Brese, & Harris 1991), three birds (Alex, Griffin, Arthur) were now together 24/7. A gradual shift was unlikely if Alex had maintained his previous pattern of vocalizing in private: significant portions of Alex's solitary practice involved what in humans would be considered rhymes (e.g. "green, cheen, bean"; "mail, banail") in which ends of labels were stable (Pepperberg, Brese & Harris 1991); that is, he seems to have (or have acquired) categorical distinctions and minimal pairs similar to those of his human models (Patterson & Pepperberg 1994, 1998). An abrupt shift could indicate some level of self-monitoring and even some additional awareness that the appropriate vowel for "spool" derived from yet another label such as "two" (/tu/); note that such information was unavailable to Arthur.

The pattern of acquisition is not unique to "spool"; I have recently found a similar pattern for Alex for the label "seven" (first in reference to the Arabic numeral, then in reference to a set of objects). Alex's initial production of the label could best be described as "s……n," a bracketing using the phonemes /s/ and /n/; he then quickly progressed to "s-one" (see **Figure 6**; /s/-pause-/wən/) which looked quite different from my "seven," but followed the form of "s-pause-wool".

After a period of two years, he replaced "s-one" with something sounding to the human ear like "seben," much closer to my "seven" (**Figure 7**; sonagraph expanded for reference).

Reprinted from LACUS Forum 34: Speech and Beyond, ed. by Patricia Sutcliffe, Lois Stanford, and Arle Lommel. Houston, TX: LACUS. 2008

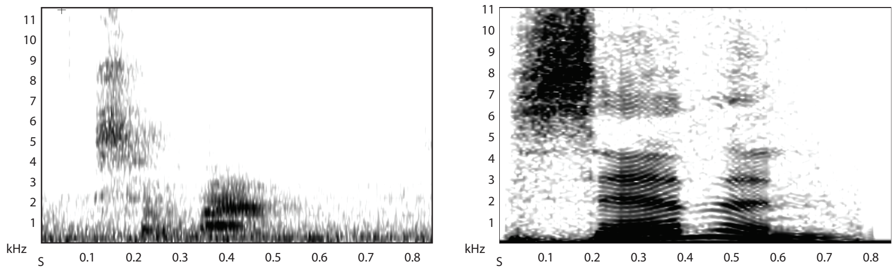

Figure 7. *(a) Alex's "seben" [/sɛbɪn/] compared to (b) Pepperberg's "seven" [/sɛvɪn/].*

3. Discussion. Alex's training likely enabled him to use phonological awareness (in the sense defined in Anthony and Francis 2005) to create a difficult new label from existing bits of sounds already in his repertoire (i.e., via segmentation) and to carefully produce the appropriate phonemes; in contrast, Arthur, who lacked such training, adapted a parrot-like whistle to produce an approximation for at least part of the novel label (Pepperberg 2007a). As expected, the parrot with the most training in vocal communication demonstrated more advanced behavior—or at least closer adherence to the performance criteria established by the human models—than did the parrot with less training. Arguably, Alex's long-term exposure to Pepperberg's speech enabled him to re-create phonetic details that were unavailable to Arthur because of the latter's relatively short exposure to human models; that is, Alex, but not Arthur, could more easily compare his output against the socially-derived human benchmark (see, e.g., Port & Leary 2005, Port 2007). Nevertheless, alternative interpretations of Alex's behavior are possible, and, unfortunately, space does not permit a full discussion here. Detailed arguments and explanations can, however, be found in Pepperberg (2007a).

I now return to the initial hypothesis, that Alex's vocal segmentation provides evidence for true imitation, rather than mimicry. Mere mimicry can be defined as purposeless duplication of an act (for a bird, rote reproduction of human speech without referential content), behavior that lacks cognitive complexity and intentionality (e.g., Tomasello & 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 presented above, Alex's data demonstrate that he has a functional understanding that his existent labels are comprised of individual units that can intentionally be recombined in novel ways to create referential, novel vocalizations (Pepperberg 2007a, b).

Although Alex's abilities are clearly not isomorphic with human language, my data (including previous studies, Pepperberg & Shive 2001, Pepperberg 2007a) demonstrate that elements of linguistic behavior, such as segmentation, are not limited to primates, nor are the neurological systems underlying such behavior. Although Alex seemingly generates novel meaningful labels from a finite set of elements, the rule system he demonstrated was relatively limited. Nevertheless, the data add another intriguing parallel between Alex's

Reprinted from LACUS Forum 34: Speech and Beyond, ed. by Patricia Sutcliffe, Lois Stanford, and Arle Lommel. Houston, TX: LACUS. 2008

and young children's early label acquisition (e.g., issues of babbling, referential and fast mapping, solitary sound play; Pepperberg 1999). And, although avian neuroanatomy and its relation to the mammalian line is not yet well enough understood to determine specific parallels among oscine, psittacine, and mammalian structures, significant progress is being made (e.g., Jarvis *et al.* 2005). Overall, despite the evolutionary distance between parrots and primates, the search for and arguments concerning responsible neural substrates and common behavior should be approached with care and not be restricted to primates (Pepperberg 2007a, b). My data, plus knowledge of avian vocal learning, of how social interaction affects such learning, and of birds' advanced cognition (e.g., Clayton *et al.* 2005, Kenward *et al.* 2005, Kroodsma & Miller 1996, Pepperberg 1999), all suggest that avian species may be important models for determining the evolutionary pressures responsible for—and in developing testable theories about—complex communication systems, particularly those involving vocal learning.

ACKNOWLEDGEMENTS

I thank Drs. Diana Reiss (Hunter College, NYC) and Donald Kroodsma (UMass-Amherst) for creating the sonagrams on *Raven* and Dr. Ofer Tchernichovski (CCNY) for the sonagram using *Sound Analysis Pro*. Research was supported by donors to *The Alex Foundation.*

REFERENCES

Anthony, Jason L. & David Francis. 2005. Development of phonological awareness. *Current directions in psychological science* 14:255–59.

Arbib, Michael. 2005. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and brain sciences* 28:105–67.

Byrne, Brian & Alvin M. Liberman. 1999. Meaninglessness, productivity and reading: Some observations about the relation between the alphabet and speech. In *Reading development and the teaching of reading*, ed. Jane Oakhill & Roger Beard, 157–74. Oxford: Blackwell.

Carroll, Julia M., Margaret J. Snowling, Jim Stevenson & Charles Hulme. 2003. The development of phonological awareness in preschool children. *Developmental psychology* 39:913–23.

Clayton, Nicola S., Joanna Dally, James Gilbert & Anthony Dickinson. 2005. Food caching by western scrub-jays (*Aphelocoma californica*) is sensitive to the conditions at recovery. *Journal of experimental psychology: Animal behavior processes* 31:115–24.

Fouts, Roger S. & Randall L. Rigby. 1977. Man-chimpanzee communication. In *How animals communicate*, ed. by Thomas Sebeok, 1034–54. Bloomington: Indiana University Press.

Reprinted from LACUS Forum 34: Speech and Beyond, ed. by Patricia Sutcliffe, Lois Stanford, and Arle Lommel. Houston, TX: LACUS. 2008

Gathercole, Susan E. & Alan D. Baddeley. 1990. The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British journal of psychology* 81:439–54.

Greenfield, Patricia M. 1991. Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and brain science* 14:531–95.

Jarvis Erich D., Onur Güntürkün, Laura Bruce, András Csillag, Harvey Karten, *et al.* 2005. Avian brains and a new understanding of vertebrate brain evolution. *Nature reviews neuroscience* 6:151–59

Kenward, Ben, Alex Weir, Christian Rutz & Alex Kacelnik. 2005. Tool manufacture by naive juvenile crows. *Nature* 433:121.

Kroodsma Donald E. & Edward H. Miller (eds.) 1996. *Ecology and evolution of acoustic communication in birds*. Ithaca NY: Cornell University Press.

Leonard, Laurence B. 2001. Fillers across languages and language abilities. *Journal of child language* 28:257–61.

Lieberman, Philip. 1984. *The biology and evolution of language*. Cambridge MA: Harvard University Press.

Mann, Virginia A. & Judith G. Foy. 2003. Phonological awareness, speech development, and letter knowledge in preschool children. *Annals of dyslexia* 53:149–73.

Patterson, Diane K. & Irene M. Pepperberg. 1994. A comparative study of human and parrot phonation: Acoustic and articulatory correlates of vowels. *Journal of the Acoustical Society of America* 96:634–48.

———. 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, Sharon. 2003. Phonological acquisition: Recent attainments and new challenges. *Language and speech* 46:87–113.

Pepperberg, Irene M. 1981. Functional vocalizations by an African Grey parrot (*Psittacus erithacus*). *Zeitschrift für Tierpsychologie* 55:391–60.

———. 1999. *The Alex studies*. Cambridge MA: Harvard.

———. 2006. Grey Parrot numerical competence: A review. *Animal cognition* 8:377–91.

———. 2007a. Grey parrots do not always "parrot": Roles of imitation and phonological awareness in the creation of new labels from existing vocalizations. *Language science* 29:1–13.

———. 2007b. When training engenders a failure to imitate in Grey parrots (*Psittacus erithacus*). In *Proceedings of the AISB '07 fourth international symposium on imitation in animals and artifacts,* ed. by Manuel Lopes. Newcastle-Upon-Tyne: University of Newcastle.

——— & Heather R. Shive. 2001. Simultaneous development of vocal and physical object combinations by a Grey Parrot (*Psittacus erithacus*): Bottle caps, lids, and labels. *Journal of comparative psychology* 115:376–84

——— & Steven R. Wilkes. 2004. Lack of referential vocal learning from LCD video by Grey Parrots *(Psittacus erithacus). Interaction studies* 5:75–97.

Reprinted from LACUS Forum 34: Speech and Beyond, ed. by Patricia Sutcliffe, Lois Stanford, and Arle Lommel. Houston, TX: LACUS. 2008

———, Katherine J. Brese & Barbara J. Harris. 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* 12:151–77.

Peters, Ann N. 2001. Filler syllables: What is their status in emerging grammar? *Journal of child language* 28:229–42.

Port, Robert. 2007. How are words stored in memory? Beyond phones and phonemes. *New ideas in psychology* 25:143–70.

Port, Robert & Adam P. Leary. 2005. Against formal phonology. *Language* 81:927–64.

Reiss, Diana & Brenda McCowan. 1993. Spontaneous vocal mimicry and production by bottlenose dolphins (*Tursiops truncatus*): Evidence for vocal learning. *Journal of comparative psychology* 107:301–12.

Studdert-Kennedy, Michael. 2002. Mirror neurons, vocal imitation, and the evolution of particulate speech. In: *Mirror neurons and the evolution of brain and language,* ed. by Maxim Stamenov & Vittorio Gallese, 207–27. Amsterdam: John Benjamins.

Thorpe, William. 1963. *Learning and instinct in animals*, 2nd ed. Cambridge MA: Harvard University Press.

Todt, Dietmar. 1975. Social learning of vocal patterns and modes of their applications in grey parrots. *Zeitschrift für Tierpsychologie* 39:178–88.

Tomasello, Michael & Malinda Carpenter. 2005. Intention reading and imitative learning. In *Perspectives on imitation: from neuroscience to social science,* ed. by Susan Hurley & Nick Chater, 135–48. Cambridge MA: MIT Press.

Treiman, Rebecca. 1995. Errors in short-term memory for speech: A developmental study. *Journal of experimental psychology: Learning, memory, and cognition* 21:1197–1208.

Vihman, Marilyn. 1996. *Phonological development: The origins of language in the child.* Malden MA: Blackwell.

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

Reprinted from LACUS Forum 34: Speech and Beyond, ed. by Patricia Sutcliffe, Lois Stanford, and Arle Lommel. Houston, TX: LACUS. 2008