

**\GREY PARROT (*PSITTACUS ERITHACUS*) 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, 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). And, of course, a robot or avatar that meaningfully imitates would circumvent many communicative challenges. Nevertheless, considerable confusion exists about the term “imitation”, which must be clarified—at least for this paper—before discussing its role in communication. The first step in such clarification is separating “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. Here I review arguments that the intentional, referential reproduction of novel English vocalizations by a Grey parrot, Alex, is imitation, particularly when the targeted novel vocalizations are constructed from related elements already in his repertoire (i.e., segmentation); I discuss consequences of imitative behavior in terms of

Alex's communicative competence (Pepperberg 2007a,b) as a model for computational studies.

Two arguments have been proposed against interpreting Alex's behavior as imitation (see Pepperberg 2007a,b). One is that avian imitation of English speech does not involve intentional, accurate reproduction of human articulatory acts. The second is that nonhumans are incapable of segmentation.

The first argument has already been countered (Pepperberg 2007a,b): Alex's parrot anatomy prevents him from exactly reproducing human articulatory acts (Patterson & Pepperberg 1994, 1998), 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 et al. 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**; Patterson & Pepperberg 1994, 1998 revealed formant structures for all Alex's vowels and stop consonants, /p,b,d,g,k,t/). Thus his speech is not simply the result of, for example, sine wave interference (Lieberman 1984).

[Insert Figure 1 here]

Countering the second argument, by claiming that Alex is capable of vocal segmentation—a special form of vocal combinatory behavior—would imply that he understands that his existent labels are formed of individual morphemes or phonemes that can be combined in novel ways to create what are for him 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 allow such combinatorial behavior (Pepperberg 2007a). Even for children, such behavior is not considered innate: 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 around three years of age or later (Carroll et al. 2003; Vihman 1996); furthermore, manipulation of individual word parts is presumed to require development of an internal representation of phonological structure (Byrne & Liberman 1999). Thus, to sound out—imitate, not 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, then recombine these elements in an appropriate sequence (Gathercole & Baddeley 1990; Treiman 1995; Arbib 2005). Moreover, children’s ability to focus on the sounds 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 nonhuman segmentation, 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 specific combinations to denote one element. Here I review evidence for Alex’s sound segmentation (Pepperberg 2007a,b, 2008).

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 (for his accomplishments, see Pepperberg 2007a); and Arthur, then only 4.5 years old. Although Arthur had had about 3.5 yrs of intense human interaction, he had had the equivalent of only about a full year of communication training (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 Model/Rival (M/R) technique (Pepperberg 1981; 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.

2. RESULTS.

Arthur's acquisition 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 consonant /l/, but had difficulty with /p/; production of a human /p/ is troublesome without lips. Unlike Alex, who learned to produce /p/ apparently via esophageal speech (Patterson & Pepperberg 1998), Arthur produced a whistled, not plosive, /p/ in /sp/ (see **Figure 2**; Pepperberg 2007a).

[Insert Figure 2 here]

His behavior suggested 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. However, only /sp/ followed the whistled pattern; /u/ (which could easily have been whistled) and /l/ resembled human speech (Pepperberg 2007a).

Alex, observing the attention that Arthur received for labeling the wooden bobbin, evinced interest in the item and 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"; /s-pause-wU/; 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 [NB: Both 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]. He knew /u/ from labels such as “two” and “blue” (Pepperberg 2007a).

[Insert Figure 3 here]

Alex retained “s-wool” for almost a year, even though usually 20-25 modeling sessions (i.e., several weeks of training) enable learning of a new label with existent phonemes (Pepperberg 1999). He then spontaneously produced a perfectly formed “spool” (/spul/). Thus, he replaced the space with /p/ and shifted the vowel toward the appropriate /u/ sound (**Figures 4,5**; Pepperberg 2007a).

[Insert Figure 4 here]

Comparing **Figures 2** and **4** shows that Arthur’s and Alex’s productions differ significantly in acoustic and sonographic patterns. Alex did not imitate or mimic Arthur. Arthur’s utterance had an avian whistle-like quality; Alex’s utterance sounded distinctly human. Alex’s utterance also resembled Pepperberg’s (**Figure 2B**), even though students had performed 90% of the training. **Figure 5** highlights how Alex’s vowel section changed from /U/ to /u/ to resemble Pepperberg’s.

[Insert Figure 5 here]

Whether Alex’s shift from /U/ to /u/ was gradual is unknown. We could not monitor his solitary practice as was done previously (Pepperberg et al. 1991) because three birds (Alex, Griffin, Arthur) were now together 24/7. A gradual shift was unlikely if Alex had maintained his previous pattern of private vocalizing: 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 et al. 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/); such information was unavailable to Arthur.

Alex’s acquisition pattern was not unique to “spool”; a similar pattern emerged for “seven” (first in reference to the Arabic numeral, then in reference to a set of objects). His initial “seven” could best be described as “s.....n”, a bracketing using the phonemes /s/ and /n/; he then progressed to “s-one” (**Figure 6**; /s/-pause-/wən/) which looked quite different from my “seven”, but followed the form of “s-pause-wool”.

[Insert Figure 6 here]

After 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).

[Insert Figure 7 here]

3. DISCUSSION.

Alex’s training likely enabled him to use phonological awareness (*sensu* Anthony & Francis 2005) to create a new label from existing bits of sounds already in his repertoire (i.e., via segmentation) and to produce the appropriate phonemes; in contrast, Arthur, lacking 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 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 (e.g., Port & Leary 2005; Port 2007). Nevertheless, alternative interpretations of Alex's behavior are possible; detailed arguments and explanations are 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 2003). 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. 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 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 and young children's early label acquisition (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).

These findings may be useful for computer scientists who are developing speech skills in atavars and robots. First, as Patterson and Pepperberg (1994, 1998) demonstrated, Alex produced his utterances with little variation in his first formant, and most of the variation in his second (and possibly third). Thus speech modeling may be simplified if based on avian, rather than human, productions. 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. 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.

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, J.L. and D. Francis. 2005. Development of phonological awareness. *Current Directions in Psychological Science* 14:255-259.
- Arbib, M. 2005. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences* 28:105-167.
- Byrne, B. and A.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*, edited by Jane Oakhill and Roger Beard, 157-174. Oxford: Blackwell.
- Carroll, J.M., M.J. Snowling, J. Stevenson and C. Hulme. 2003. The development of phonological awareness in preschool children. *Developmental Psychology* 39:913-923.
- Fouts, R.S. and R.L. Rigby. 1977. Man-chimpanzee communication. In *How animals communicate*, edited by Thomas Sebeok, 1034-1054. Bloomington: Indiana University Press.
- Gathercole, S.E. and A.D. Baddeley. 1990. The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology* 81:439-454.
- Greenfield, P.M. 1991. Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavior and Brain Science* 14:531-595.
- Jarvis .D., O. Güntürkün, L. Bruce, A. Csillag, H. Karten et al. 2005. Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience* 6:151-159
- Leonard, L.B. 2001. Fillers across languages and language abilities. *Journal of Child Language* 28:257-261.
- Liberman, P. 1984. *The biology and evolution of language*. Cambridge, MA: Harvard

University Press.

- Mann, V.A. and J.G. Foy. 2003. Phonological awareness, speech development, and letter knowledge in preschool children. *Annals of Dyslexia* 53:149-173.
- Patterson, D.K. and I.M. Pepperberg. 1994. A comparative study of human and parrot phonation: acoustic and articulatory correlates of vowels. *Journal of the Acoustical Society of America* 96:634-648.
- . 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 and Speech* 46:87-113.
- Pepperberg, I.M. 1981. Functional vocalizations by an African Grey parrot (*Psittacus erithacus*). *Zeitschrift für Tierpsychologie* 55:39-160.
- 1999. *The Alex studies*. Cambridge MA: Harvard.
- . 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*, edited by Manuel Lopes, Newcastle-Upon-Tyne, UK: University of Newcastle.
- and H.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* 115:376-384
- and S.R. Wilkes. 2004. Lack of referential vocal learning from LCD video by Grey Parrots (*Psittacus erithacus*). *Interaction Studies* 5:75-97.

- , K.J. Brese and B.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* 12:151-177.
- Peters, A.N. 2001. Filler syllables: What is their status in emerging grammar? *Journal of Child Language* 28:229-242.
- Port, R. 2007. How are words stored in memory? Beyond phones and phonemes. *New Ideas in Psychology* 25:143-170.
- Port, R. and A.P. Leary. 2005. Against formal phonology. *Language* 81:927-964.
- Reiss, D. and B. McCowan. 1993. Spontaneous vocal mimicry and production by bottlenose dolphins (*Tursiops truncatus*): Evidence for vocal learning. *Journal of Comparative Psychology* 107:301-312.
- Studdert-Kennedy, M. 2002. Mirror neurons, vocal imitation, and the evolution of particulate speech. In: *Mirror neurons and the evolution of brain and language*, edited by Maxim Stamenov and Vittorio Gallese, 207-27. Amsterdam: John Benjamins.
- Thorpe, W. 1963. *Learning and instinct in animals*, 2nd edited Cambridge, MA: Harvard University Press.
- Todt, D. 1975. Social learning of vocal patterns and modes of their applications in Grey parrots. *Zeitschrift für Tierpsychologie* 39:178-188.
- Tomasello, M and M. Carpenter. 2005. Intention reading and imitative learning. In *Perspectives on imitation: from neuroscience to social science*, edited by Susan Hurley and Nick Chater, 135-148. Cambridge, MA: MIT Press.
- Treiman, R. 1995. Errors in short-term memory for speech: A developmental study. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21:1197-1208.

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

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

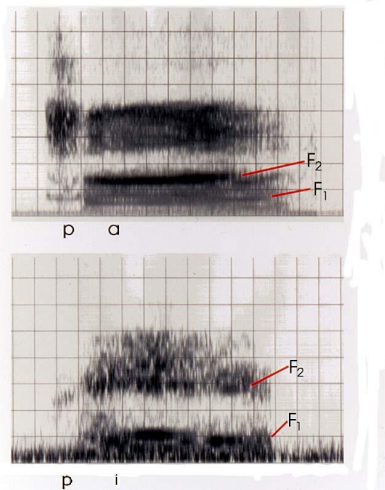


Figure 1: Alex's production of "pah" /pa/, his label for pasta and for "pea" /pi/, his label for a green pea.

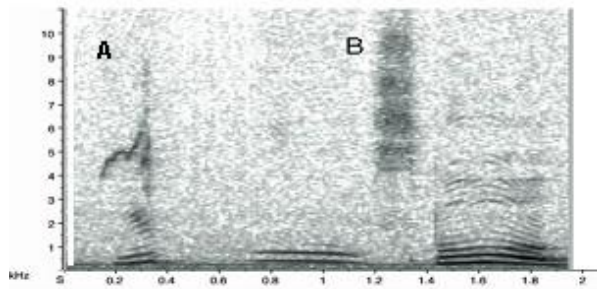


Figure. 2: (a) Arthur's "spool" compared to (b) Pepperberg's "spool" (from Pepperberg 2007a)

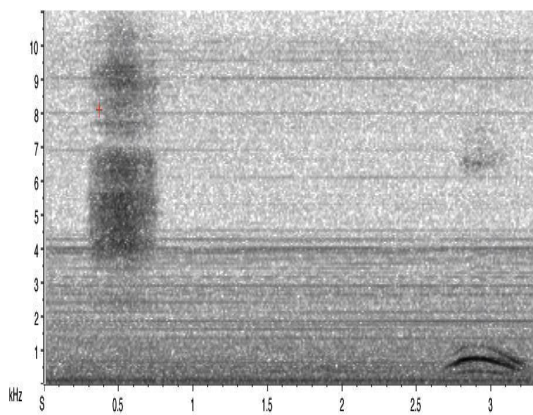


Figure 3: Alex's "s-wool" (/s-pause-wUl/), (from Pepperberg 2007a)

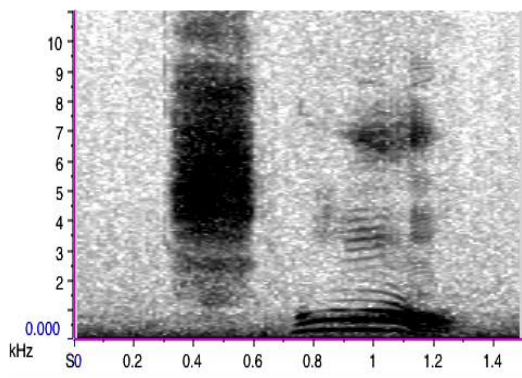


Figure 4: Alex's "spool" (/spul/) (from Pepperberg 2007a)

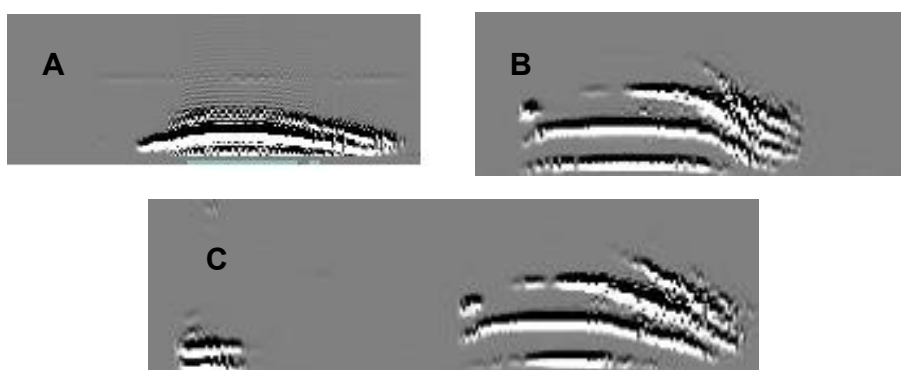


Figure 5: (a) Alex's /U/; (b) Alex's /u/, (c) part of Pepperberg's /spu/ (from Pepperberg 2007a)

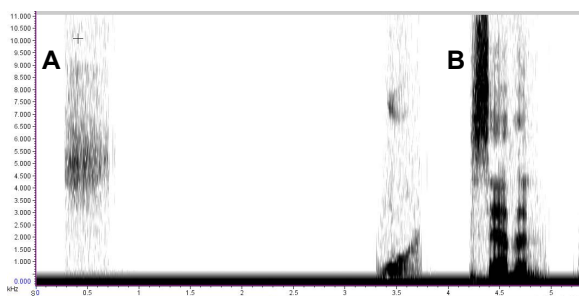


Figure 6: (a) Alex's "s-one" /s-pause-wən/ followed by (b) Pepperberg's "seven" /sEvIn/

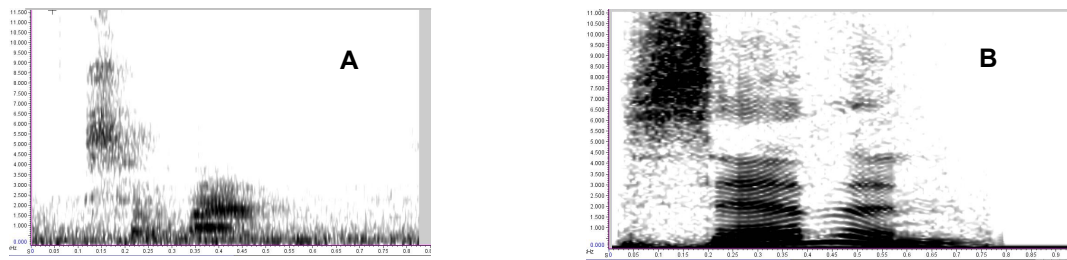


Figure 7: (a) Alex's "seben" [sEbIn/] compared to (b) Pepperberg's "seven" [sEvIn/]