

A FIELD TEST OF SYRINX SOUND ANALYSIS SOFTWARE IN INTERACTIVE PLAYBACK

DANIEL J. MENNILL¹ AND LAURENE M. RATCLIFFE

Department of Biology, Queen's University, Kingston, Ontario, K7L 3N6, Canada

ABSTRACT

Syrinx is a newly developed PC-based software package designed for interactive playback of animal vocalizations and for digital sound recording and analysis in the lab or in the field. Syrinx permits flexible modeling of complex vocal interactions in an interactive playback setting. Large numbers of sounds can be stored on disk as digital wave files for instant recall in the field. The playback trial timer and real-time scrolling spectrographic display of microphone input facilitate the tailoring of playback responses to subtle changes in the vocalizations of playback subjects. Based on our field test of this software in an interactive playback experiment with Black-capped Chickadees (*Poecile atricapillus*), we conclude that Syrinx is a versatile and user-friendly tool for performing interactive playback with a minimum of equipment.

INTRODUCTION

Traditional non-interactive playback has been a powerful tool for experiments designed to understand fundamental questions in animal communication (McGregor 1992). However, its utility for investigating complex, dyadic interchange between senders and receivers is limited. For example, investigators using non-interactive approaches employing pre-recorded tape loops cannot select one of a catalogue of stimuli quickly or easily for playback in the field. Hence non-interactive playback cannot easily mimic animals which produce multiple vocalization types or otherwise variable vocalizations. Furthermore, non-interactive techniques such as loop playback do not allow the investigator to tightly regulate the timing of stimuli given in playback and therefore cannot be used to simulate vocal interactions where the relative timing of signals is important. Interactive playback is an emerging technique which attempts to overcome these limitations by allowing the investigator to instantaneously choose from a variety of stimuli and regulate the timing of

¹Corresponding Author. E-mail: mennilld @ biology.queensu.ca

playback of these stimuli (Dabelsteen & McGregor 1996). Several investigations demonstrate that interactive playback is important for the presentation of relevant and meaningful stimuli in the study of the function of animal vocalizations (e.g. Nielsen & Vehrencamp 1995, Smith 1996, Otter et al. 1999).

Digital technology allows the storage and rapid recall of sampled sounds and thus the capacity for interactive playback. Dabelsteen and Pedersen's (1991) digital sound emitter (DSE) was the first playback system designed to take advantage of storage and recall of digitally sampled sounds in the field. Recently, with the development of lightweight, high-speed, high-storage-capacity laptop computers, two new interactive playback systems for the Macintosh emerged; Bradbury and Vehrencamp's (1995) Singit! software and Otter et al.'s (1995) HyperCard stack technique both permit interactive playback using multiple sounds stored on a laptop computer's hard drive. All three of these systems, however, rely on additional software or hardware for acquisition of stimuli, and none can be used to edit sounds in the field. Furthermore, these systems lack a method for discriminating between playback subjects' vocalizations (such as a real-time spectrographic display) and therefore rely heavily on the investigator's discrimination ability.

In this paper we review Syrinx sound analysis software, a new PC-based program designed with interactive playback specifically in mind. We describe Syrinx's principal recording features, sound analysis features and interactive playback features. We outline our field test of Syrinx in an interactive playback experiment with Black-capped Chickadees (*Poecile atricapillus*) to demonstrate some of the possibilities of this software. In our opinion, Syrinx is an important, accessible and versatile tool which presents tremendous opportunities for modeling complex vocal interactions through interactive playback.

Technical requirements

Syrinx runs in Windows 95 and 98 based systems and is presently being programmed for Windows NT compatibility (John Burt pers. comm.). A Macintosh version is not available. Researchers using the scrolling spectrographic display feature of Syrinx require a minimum CPU speed of 133 MHz, although 266 MHz (or better) is recommended. Syrinx can be run on slower systems providing the scrolling spectrographic display is disabled. Syrinx requires an 8 bit (or higher) soundcard and both a microphone input jack and a speaker or headphone output jack. A full duplex sound card is not required but will allow playback without interruption to the scrolling spectrographic display (John Burt pers. comm.).

For our interactive playback field test of Syrinx, we selected a Panasonic Toughbook laptop computer (Model CF-45; Pentium MMX, 266 MHz, 32 KB RAM, 16 bit sound card) because of its durability and long battery life. We used a Sennheiser directional microphone (Model MKH 816) for live

input into Syrnix and a Sony active speaker (Model SRS 77G) for playback. Turning off the microphone volume in the computer's audio output settings prevented feedback between the microphone and speaker. We ran Syrnix successfully in both Windows 95 and Windows 98 platforms on this computer.

Display and recording features of Syrnix

The visual display of Syrnix is established by the user. It consists of multiple resizable windows each containing a sound spectrogram or waveform amplitude display of a digitally stored file (Figure 1). For each window on the display the

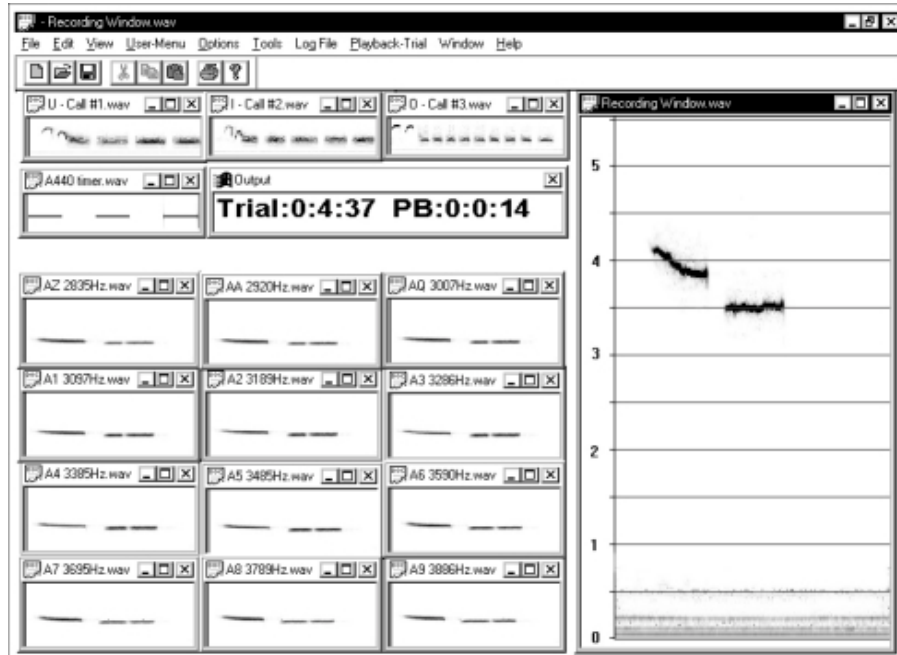


Figure 1. An example of the Syrnix screen display. In this figure, Syrnix is designed to simulate a Black-capped Chickadee capable of frequency shifting and frequency matching the songs of playback subjects. The large window at right is a real-time scrolling spectrograph used to assess the absolute frequency of the playback subject's songs as they are recorded from a microphone input directly to disk. The output window shows the time since the beginning of the playback trial and the time since the most recent playback. The remaining 16 windows are sound spectrograms of sound files stored on the computer's hard disk (including three *chickadee* calls at the top left and 12 chickadee songs of different absolute frequency at the bottom left) and used as playback stimuli in our interactive experiment.

user may adjust the FFT resolution (32 to 1024 Hz), the FFT function (Blackman, Hamming or Hanning), the spectral gain and white threshold, and the time and frequency axis features of spectrograms. (As such, Syrinx is a useful tool for generating high-quality graphics for use in displays and presentations either by printing spectrograms directly from Syrinx or by exporting bitmap graphics to other software packages.) Users can store multiple display formats and display settings to use for different applications of Syrinx. For example, we used one display format for recording and analyzing sounds and another for interactive playback.

During recording with Syrinx, either from a live source or from an analogue or digital recording, an active window shows a real-time scrolling spectrographic display and an amplitude display of the input simultaneously. Sounds can be stored directly to disk as digital wave files at a sampling rate of 8800 to 44100 Hz (wave file size = 44.1 KB/second for mono, 16-bit recording at a sampling rate of 22050 Hz; not all sound cards can sample at all sampling rates). The real-time scrolling spectrographic display can be assigned an FFT resolution of 32 to 1024 Hz and a Blackman, Hamming or Hanning window transformation. While recording, the user may visually scan the scrolling display for desired sounds and store the buffered sounds to disk with a single keystroke, at which point the scrolling display continues in a new window. Users may customize Syrinx to automatically filter, amplify, or trim buffered sounds as they are saved to disk during recording.

The Automated Sound Detector is, by far, the most innovative aspect of Syrinx's recording features, allowing a computer to act as a "smart" voice-activated recorder. The user dictates what sound events will trigger the Automated Sound Detector by specifying the trigger threshold above noise level (%), the minimum total sound duration and minimum trigger sound duration (microseconds), the maximum gap between trigger events (microseconds) and the minimum and maximum frequency range of the trigger event (Hz). Since Syrinx stores input to a temporary buffer there is no cropping of recorded sounds due to delay between signal input and trigger event. We used the Automated Sound Detector to isolate Black-capped Chickadee and Yellow Warbler (*Dendroica petechia*) songs from long analogue field recordings; after fine-tuning the trigger event settings for each species, Syrinx reliably retrieved all songs from six 90-minute tapes and stored them using a minimum of drive space. The automated detector would also be useful for recording animals in laboratory experiments where continuous monitoring of vocalizations is required.

Sound analysis and manipulation features

One of the major advantages of Syrinx compared to other interactive playback software is that Syrinx integrates playback features with basic sound-analysis and sound-editing features. As such, the user can evaluate and alter sounds quickly in the field without relying on other software packages. Stored sounds are analyzed and edited as spectrographs. This permits a high level of precision while taking measurements using time and frequency cursors, trimming unwanted sections of a sound file, or adding noise, tones or silences to existing sound files. Time and frequency measurements made with the time and frequency cursors can be stored in a tab-delimited log file for export as a database. High-pass, low-pass and band-reject filtering may be performed using the cursors or by assigning preset filter values. Syrinx can also be used to apply amplitude envelopes and amplitude ramps, equalize sounds (by amplifying the loudest part of a sound file to peak amplitude) or reduce background noise (by calculating a background noise profile and increasing the signal to threshold level using this profile).

Interactive playback features

In studies of songtype matching and song overlapping in Banded Wrens (*Thryothorus pleurostictus*; Molles and Vehrencamp unpubl.), Great Tits (*Parus major*; Otter et al. 1999) and Song Sparrows (*Melospiza melodia*; Nielsen and Vehrencamp 1995, Burt, Campbell and Beecher unpubl.) investigators have designed interactive playback sessions so that each playback response varies according to each vocalization given by the playback subject; that is, the playback subject, not the investigator, drives the interaction. Like other interactive playback techniques (Dabelsteen and Pederson 1991, Bradbury and Vehrencamp 1994, Otter et al. 1994), Syrinx permits subject-driven interactions by allowing the investigator to reply to subjects' vocalizations using multiple playback stimuli from active on-screen windows. On-screen stimuli can be selected immediately with a mouse click or by assigning a keystroke to each stimulus (less than 0.25 seconds delay between a keystroke and sound activation when scrolling spectrographic display is active; 0.0 seconds delay between keystroke and sound activation when scrolling spectrographic display is not active). Using Syrinx's log file an investigator can keep a record of all stimuli given in playback, assign log file hot keys to quickly mark events in the log file during playback (e.g. specific behaviours of playback subjects) or enter comments into the log file to produce detailed accounts of playback trials.

Previous interactive playback techniques have been limited by the number of stimuli that can be stored on the interactive playback device, by the investigator's estimates of the timing of their playback responses or by the investigator's ability to recognize song types and choose an appropriate playback response (Dabelsteen and McGregor 1996). Syrinx overcomes all

three of these limitations. The number of sounds that may be stored on-screen for instant recall is limited only by the size of the screen (on our 12-inch screen at a resolution of 800 by 600 pixels we could discriminate up to 30 spectrograms, although discrimination ability will vary with the structural characteristics of the sounds of interest). Secondly, using Syrinx's on-screen playback trial clock (see Figure 1) in conjunction with the time axis on the scrolling spectrographic display, playback responses can be timed more precisely than by ear or stopwatch alone. Finally, the real-time scrolling spectrograph of Syrinx allows the experimenter to match songtypes precisely by visually matching on-screen spectrograms of stored stimuli to the real-time spectrogram of the playback subject's songs. There are inherent limits to the human ability to discern differences in vocalizations of many animal species (e.g. species with large song repertoires such as mockingbirds or nightingales, species with very subtle distinctions between vocalizations such as frequency variation in the songs of Black-capped Chickadees, or species which vocalize outside the range of human hearing such as bats). However, the match-and-playback design of Syrinx may allow interactive playback experiments involving species with vocalizations too diverse or subtly different to discern otherwise.

An interactive playback field test of Syrinx on Black-capped Chickadees

In spring 1999 we used Syrinx in an interactive playback experiment as part of our ongoing research of communication networks and eavesdropping in Black-capped Chickadees at Queen's University Biological Station near Kingston, Ontario, Canada. The Black-capped Chickadee is a resident songbird which sings a single, two-note songtype often designated "*fee-bee*". Although Black-capped Chickadees have only one songtype they can shift the absolute frequency of their songs across a range of approximately 2900 Hz to 3500 Hz (as measured by the frequency of the *bee* note; Horn et al. 1992). In countersinging interactions between males during the breeding season, individuals sometimes match the frequency and overlap the songs of rivals. Both of these signals are thought to indicate directed aggression between countersinging chickadees (Shackleton and Ratcliffe 1994). In our playback experiment we tested the behavioural implications of frequency matching and song overlapping using interactive playback. Here we describe our experimental method and some preliminary results to highlight the capabilities of Syrinx.

To create playback stimuli we used Syrinx to digitize six songs from six different chickadees recorded in 1987. We isolated the *fee* note from three birds' songs and the *bee* note from the other three birds' songs and combined them using Syrinx's trimming and pasting features to create three population-typical chickadee songs. Using CoolEdit 96 software (Syntrillium Group 1996) we transposed these three songs across the species frequency range at intervals of 100 Hz to create three song libraries. We used one of the three sound

libraries in each playback trial to simulate a chickadee capable of matching the frequency of any playback subject's song to within 50 Hz. To avoid pseudoreplication we used a latin square to randomly choose one of the three song libraries for playback in 32 trials to 32 different birds.

Playback trials began with a timer beep (to synchronize Syrinx's log file with an observing researcher's tape recorder) which activated both Syrinx's playback trial clock as well as a running log file. We then used tandem repeats of "*chickadee*" calls for a maximum of 3.0 minutes to lure territorial males to the speaker location. As soon as the focal male responded to the lure by singing, we rapidly assessed the absolute frequency of the playback subject's song from Syrinx's real-time scrolling spectrographic display and used this information to gauge our response (see below). Playback trials lasted 6.0 minutes or until the focal male stopped singing.

We used two different interactive playback protocols to respond to focal males (Figure 2). In one protocol we answered every song given by the focal male by responding within 1.0 seconds of the beginning of the focal male's song and by frequency matching the focal male's song to within 50 Hz. If the focal male shifted the frequency of his song we immediately shifted the frequency of the playback response to match him. In our second protocol we delayed our playback responses (i.e. 2.0 to 2.5 seconds between the beginning of the focal male's song and the beginning of the playback response) and avoided frequency matching by offering playback responses 200 to 300 Hz higher than the focal male's songs. We did not alter the frequency of our playback responses under this second protocol unless the focal male shifted the frequency of his song to match the frequency of the previous playback response, in which case we immediately shifted the frequency of playback responses to be 200 to 300 Hz higher than the focal male's new frequency.

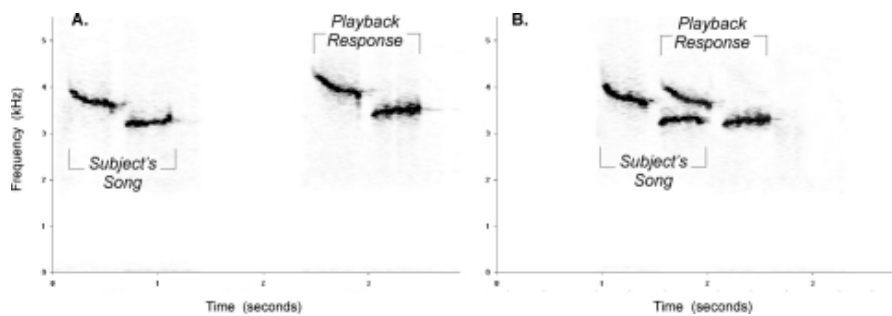


Figure 2. Sound spectrograms of two interactive playback exchanges. In one interaction protocol (spectrogram A at left), we responded to the playback subject's song by avoiding frequency matching (i.e. responding with a song 200 to 300 Hz higher than the frequency of the subject's song) and by delaying the playback response by 2.0 to 2.5 seconds after the beginning of the subject's song. In our second interaction protocol (spectrogram B, at right), we responded to the playback subject's song by frequency matching (with an accuracy of 50 Hz) and by replying within 1.0 seconds from the beginning of the subject's song. Spectrograms shown here were generated with Syrinx.

Using the scrolling spectrographic display, we were able to frequency match and reply directly to all focal males' songs ($n = 16$ six-minute frequency matching interactions) with < 0.2 seconds latency and with higher accuracy than by responding by ear alone (Otter et al. 1994). Responses of males to both playback protocols were strong; the average distance of closest approach to the speaker was 2.75 ± 0.59 (SE) meters for direct-response, frequency matching trials and 3.42 ± 0.52 (SE) meters for delayed response, matching-avoidance trials. By comparison, Fotheringham and Ratcliffe (1995) found a much higher average distance of closest approach of > 9.0 meters using non-interactive loop playback ($n = 20$ playback trials to Black-capped Chickadees). These results, taken together with other interactive playback studies, attest to the effectiveness of interactive playback for engaging playback subjects in complex vocal interactions.

Univariate analysis of male responses to the two interaction protocols shows a non-significant difference between playback subjects' distance of closest approach (two-tailed $t_{0.05,30} = 0.914$, $P = 0.37$, $n = 16$ trials for each protocol). However, after comparing nine intensity-of-response variables (distance of closest approach, latency from beginning of interaction to closest approach, number of passes over the speaker, total number of songs given by focal male, proportion of total number of songs given within 10 m of the speaker, length of interaction, proportion of interaction spent within 1 m of the speaker, between 1 and 5 m, and between 5 and 10 m) there is evidence of a differential response to the two treatments (8 of 9 variables where playback subjects found the direct-response, frequency matching protocol more aversive; 1-tailed sign test, $P = 0.02$).

Several features of our interactive playback experiment would not have been possible without Syrinx. Firstly, the real-time scrolling spectrographic display allowed us to assess the specific characteristics of playback subjects' songs as they were sung and to then choose responses rapidly and quantitatively. While Black-capped Chickadees possess only one songtype, our frequency matching experiment is analogous to a songtype-matching experiment where the investigator would read Syrinx's real-time spectrographic display to assess the playback subject's songtype and choose an appropriate response. Secondly, the scrolling spectrographic display in conjunction with Syrinx's playback trial clock allowed us to time our playback responses very precisely. Finally, Syrinx allowed the storage and rapid recall of a relatively large number of stimuli as well as the storage of playback subject's songs and log file records during playback trials.

CONCLUDING REMARKS

In our field test of Syrinx, the software proved to be a very useful tool for engaging playback subjects and modeling complex vocal exchanges through interactive playback. Using a laptop computer, a microphone and a speaker

(relatively inexpensive, transportable equipment which many investigators already possess), Syrinx will allow investigators to better study animal vocal communication in the field. We recommend this software highly and feel that Syrinx, as an integrated software package, will play an important role in many future interactive playback experiments. In particular, the Automated Sound Detector feature for sound recording and the real-time scrolling spectrographic display feature for interactive playback set Syrinx apart from other software packages.

Syrinx was written in C++ by John Burt of the University of Washington. Although Syrinx is still undergoing changes, the program is fully functional for all of the features described here. The only limitations we found with Syrinx are features already targeted by the programmer for improvement, including scroll bars within each spectrographic window, increased printing capabilities and the ability to work in multiple colour modes (Syrinx presently works only in 256 colour mode; John Burt pers. comm.). Planned additions to Syrinx include the ability to playback from lists of sounds in each window (instead of just one sound per window), the ability to create customized macros to handle complex playback contingencies and possibly the addition of a neural net for input recognition (John Burt pers. comm.). Syrinx is presently available at \$300 US to interested researchers who should visit the Syrinx website (zipprong.psych.washington.edu/syrinx.html) or contact John Burt by e-mail ([quill @ bigfoot.com](mailto:quill@bigfoot.com)). Site licenses and customized packages of Syrinx are also available.

ACKNOWLEDGEMENTS

We would like to thank A. MacDougall for field assistance with playback trials and the staff at Queen's University Biological Station for logistical support. M. Naguib, P. Christie and an anonymous reviewer provided helpful comments on the manuscript. J. Burt checked the manuscript for technical accuracy. Funding for this project was provided by an NSERC research grant to L. M. Ratcliffe and a Queen's Graduate Fellowship, an NSERC post-graduate scholarship, an Animal Behaviour Society Research Grant, a Frank M. Chapman Memorial Fund Award and an E. A. Bergstrom Memorial Research Award to D. J. Mennill.

REFERENCES

- Bradbury, J. W. & Vehrencamp, S. L. (1994). Singit!: A program for interactive playback on the Macintosh. *Bioacoustics*, **5**, 308 – 310.
- Dabelsteen, T. & McGregor, P. K. (1996). Dynamic Acoustic Communication and Interactive Playback. In *Ecology and Evolution of Acoustic Communication in Birds* (D. Kroodsma & E. Miller, eds.). Cornell University Press; Ithaca, pp. 398 – 408.
- Dabelsteen, T. & Pedersen, S. B. (1991). A portable digital sound emitter for interactive playback of animal vocalisations. *Bioacoustics*, **3**, 193 – 206.

- Fotheringham, J., & Ratcliffe, L. (1995). Song degradation and estimation of acoustic distance in Black-capped Chickadees (*Parus atricapillus*). *Can. J. Zool.*, **73**, 858 – 868.
- Horn, A. G., Leonard, M. L., Ratcliffe, L., Shackleton, S., & Weisman, R. G. (1992). Frequency variation in the songs of black-capped chickadees (*Parus atricapillus*). *Auk*, **109**, 847 – 852.
- McGregor, P. K. (ed.) (1992). *Playback and Studies of Animal Communication*. New York: Plenum Press.
- Nielsen, B. M. B. & Vehrencamp, S. L. (1995). Responses of song sparrows to song-type matching via interactive playback. *Behav. Ecol. Sociobiol.*, **37**, 109-117.
- Otter, K., McGregor, P. K., Terry, A. M. R., Burford, F. R. L., Peake, T. M., & Dabelsteen, T. (1999). Do female great tits (*Parus major*) assess males by eavesdropping? A field study using interactive song playback. *Proc. R. Soc. Lond. B*, **266**, 1305-1309.
- Otter, K., Njegovan, M., Naugler, C., Fotheringham, J., & Ratcliffe, L. (1994). A simple technique for interactive playback experiments using a Macintosh Powerbook computer. *Bioacoustics*, **5**, 303-308.
- Shackleton, S. A. & Ratcliffe, L. (1994). Matched counter-singing signals escalation of aggression in Black-capped Chickadees (*Parus atricapillus*). *Ethology*, **97**, 310 – 316.
- Smith, W. J. (1996). Using interactive playback to study how songs and singing contribute to communication about behaviour. In *Ecology and Evolution of Acoustic Communication in Birds* (D. Kroodsma & E. Miller, eds.). Cornell University Press; Ithaca, pp. 377 – 397.

Received 17 September 1999, revised 14 January 2000 and accepted 20 January 2000