THRESHOLDS OF AUDIBILITY FOR BONE-CONDUCTION HEADSETS

Bruce N. Walker and Raymond M. Stanley

Sonification Lab,
School of Psychology
Georgia Institute of Technology
654 Cherry Street
Atlanta, Georgia USA 30332.
bruce.walker@psych.gatech.edu, gtg141n@mail.gatech.edu

ABSTRACT

Despite advantages of using headphones, including privacy and portability, headphones have one essential drawback: they cover the ears of the listener, thus deteriorating detection and localization of ambient sounds. Bone-conduction headsets leave the ears uncovered, yet maintain portability and privacy. An initial step in establishing guidelines for using these "bonephones" is taken in the present research. The input into the bonephones necessary to reach a 71% detection threshold is measured at critical band centers ranging from 150 Hz to 13500 Hz. These thresholds were measured with an open ear canal, a plugged ear canal, and a masking noise. Results were consistent with other bone-conduction threshold measurements. The utility of this information in the context of equalization for the audio presented through the bonephones is discussed.

1. INTRODUCTION

1.1. The Need for Headphone Alternatives

Headphones and most other audio output hardware deliver acoustic signals to the cochlea through the medium of air, so sounds and hearing are usually discussed in terms of pressure waves transmitted through air. However, the auditory system is also sensitive to pressure waves transmitted through the bones in our skull [1, 2].

Although hearing via bone-conduction occurs naturally in listening to one's own voice and to loud external sounds, sound can also be directly transmitted through the bone via vibrators attached to the skull. Presenting auditory information to listeners through bone-conduction by placing vibrators on the skull avoids covering the ears of the listener, unlike standard headphones. Covering the ears deteriorates detection and localization of ambient sounds that may be of particular interest in tactical situations, in augmented reality applications, and for visually impaired users who rely on spatial audio cues as their primary sense of orientation (e.g., [3]). Furthermore, most headphones do not allow auditory display to occur in conjunction with hearing protection. Using bone-conduction devices matches the privacy and portability that headphones offer, yet leaves the ear canal and pinna uncovered. This may facilitate improvement in the detection and localization of environmental sounds, and allows the display of auditory information with hearing protection inserted into the ear canal.

Recently, binaural bone-conduction headsets have become available. Due to their comfort, small size, and standardized input jack, these newer "bonephones" are now suitable for implementation in auditory displays. The transducers of the bonephones rest on the mastoid, which is the raised portion of the temporal bone located directly behind the ear. The mastoid is a preferable transducer location relative to the forehead or temple because it contains the inner ear, is relatively immune to the interference associated with muscle tissue operating the jaw, and allows stereo presentation of sounds.

1.2. The Need For Bone-conduction Research

Most of psychoacoustics research and all of the human factors research on auditory displays that the authors are aware of has focused on the conduction of sound through air, and thus has overlooked the alternative acoustic pathway of bone-conduction. Since guidelines established for air conduction will not necessarily apply to bone-conduction, effective auditory display design needs to be re-evaluated for bone-conduction. Basic psychophysical data such as audibility thresholds pave the way and constrain the solution space for designers building auditory displays. For example, if a listener is unable hear a sound, he or she cannot extract information from the sound to perform the listening task. On the other hand, a sound that is too loud may block out other important auditory information, or may even damage hearing.

1.3. Bone-conduction Literature

The research that has been done on bone-conduction is limited in its applicability to the goal of understanding the basic psychological and acoustical parameters of binaural bone-conduction headsets developed for non-clinical purposes. The goals of that research have been to establish threshold norms for clinical testing of middle ear disorders or to tease apart mechanisms underlying hearing via bone-conduction. This is typically done with an apparatus consisting of a single mechanically-driven vibrator (e.g., the Radioear B-71) placed on the forehead or the mastoid. Bone-conduction threshold measurements are highly variable; this variability is due to vibrator design and placement, whether the ear canals are occluded, as well as methods of testing and measurement [1, 4, 5]. Despite the fact that the goals and apparatus differ significantly from ours, clinical research does provide estimates of the audibility threshold curve for the bonephones, as well as some insight into testing and measuring bone-conduction thresholds.

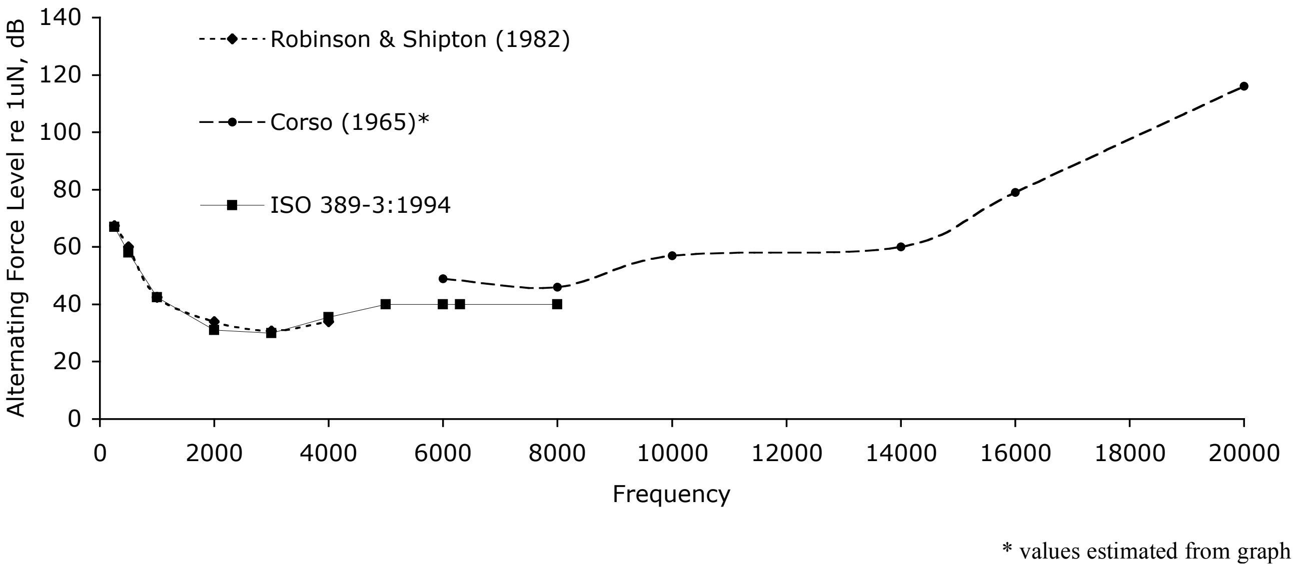

Figure 1: *Some bone-conduction thresholds as reported in the literature.*

The most up-to-date and aggressively researched thresholds are those standardized for classification of middle ear functionality (i.e., [6]). Standardized threshold values are measured with a single vibrator applied to one mastoid while the contralateral ear is masked, to prevent air or bone-conducted signals leaking to the other ear. Sometimes one (the tested) or both ears are plugged at the higher frequencies, to prevent leakage of airborne sounds. The unit of measurement is force, calibrated to a standardized artificial mastoid. Some typical bone-conduction thresholds can be found in Figure 1. The most recent public research done on standardized bone-conduction thresholds for normal adult listeners was done by Robinson and Shipton [7]; their thresholds are also shown in Figure 1. Thresholds shown are tested under conditions similar to those for ISO standardization. The range of frequencies tested in bone-conduction thresholds is often not very large, due to the clinical motivations of assessing ability to hear sounds that are most important and common in the environment. In the context of an auditory display, however, the full range of frequencies to which humans are sensitive may be useful. Although testing of higher frequencies with bone-conduction is rare and unstandardized, Corso [8] measured bone-conduction thresholds at higher frequencies with procedures similar to the ISO standards, but did not have any masking stimuli; his values can also be seen in Figure 1. Furthermore, bone-conduction may facilitate sensitivity to sounds even higher in frequency than can be detected through air-conduction, possibly up to 95 kHz [8].

Although most threshold assessments always plug one or both the ears, research has also been conducted that shows the effect of plugging the ears. With a vibrator placed on the forehead, Watson [9] compared thresholds with binaurally plugged ears and unplugged ears at frequencies in the range of 62.5 to 2000 Hz. He found the plugged threshold is 10-20 dB lower than the unplugged threshold at low frequencies (200 Hz and below), begins to converge with the unplugged threshold at medium frequencies (250 – 1000 Hz), and the two are equal above 2000 Hz. This decrease in threshold (with plugged ears) at lower frequencies is known as the "occlusion effect" [10]. Bekesy [1] also found that, when a vibrator was applied to the forehead, bone-conduction thresholds with the ears plugged were consistently lower than with the ears open, up to about 2000 Hz, after which thresholds for open and plugged ears became nearly identical.

Listening conditions such as masking and plugging the ear canal are used in an attempt to control leaking air-conducted sounds and bone-conducted sounds traveling across the skull. However, they introduce their own set of confounds and thus do not necessarily tease apart bone versus air conduction [2]. Although some aspects of the proximal listening environment have been standardized in recent years, comparing threshold values from different studies can be difficult due to variation in the listening conditions. Despite these shortcomings, these listening conditions have practical relevance in the real world, especially in the context of a mobile device displaying auditory information. Specifically, masking provides a more ecologically valid environment than a sound booth by simulating noise that would occur naturally in the environment. In addition, earplugs are often used for hearing protection in settings where auditory displays would be useful. The high variability associated with use of different vibrators and listening conditions, as well as the sparse assessment of binaural bone-conduction, warrants an in-depth look at thresholds with modern bonephones.

2. METHODS

2.1. Participants

Five graduate students at Georgia Institute of Technology (3 males, 2 females, mean age = 24.6, range = 23-26) participated. Participants received $10 for each of three sessions. Participants were screened in each session for normal hearing with a Micro Audiometrics Corporation audiometer. "Normal hearing" was defined as sensitivity to a 20 dB SPL pure tone with frequencies of 250, 500, 750, 1000, 1500, 2000, 3000, 4000, 6000, and 8000 Hz.

2.2. Stimuli

Participants listened to 1-second long binaurally presented pure tones (.wav format, 44.1 kHz sampling rate, 16-bit depth). Pure tones were generated for each of the following critical band centers: 150, 250, 350, 450, 570, 700, 840, 1000, 1170, 1370, 1600, 1850, 2150, 2500, 2900, 3400, 4000, 4800, 5800, 7000, 8500, 10500, 13500 (see [11]). The critical band center of 50 Hz was not tested because the bonephones did not have the capability to accurately and reliably reproduce this frequency. For each frequency, a set of tones was generated with attenuation levels ranging from 0 to -135 dB, in 3dB steps.

The pink noise for the masking condition was digitally generated (in .wav format) and written onto a CD with sampling rate of 44.1 kHz and bit depth of 16. The masking was played at a level of 45 dBA SPL, as measured at the approximate center of the listeners' heads by an Extech Instruments 407750 Digital Sound Level Meter.

2.3. Apparatus

Visual information was presented on a 14-inch (35.56 cm) Viewsonic LCD monitor, and participants made their responses using a standard computer keyboard. The presentation of the stimuli was controlled with a program written in Eprime (Psychology Software Tools, Inc.), running on a Dell Dimension 8100 Pentium-4 PC with Windows XP. The digital sound files were output via a Creative Labs Audigy sound card, and then amplified with a Behringer HA4600 headphone amplifier, to which the bonephones were connected. Participants listened to tones delivered to their mastoid with Temco bone-conduction headsets. Under the masked listening condition, the pink noise was played from a Samsung DVD-V1000 player, amplified with a Denon DRA-275R stereo receiver, and delivered through Klipsch KSB 1.1 speakers. The speakers were located directly out to each side of the listener (i.e., at –90 and +90 degrees), approximately 42 inches (106.68 cm) from the center of the listener's head. Under the plugged listening condition, participants inserted E·A·R foam earplugs into their ear canals. Participants completed the procedure in an Industrial Acoustics Company sound-attenuated room, with the computer located outside of the room.

2.4. Procedure

Each participant completed three sessions lasting approximately two hours each. Sessions were separated by one day of rest and each involved a different listening condition (open, plugged, or masked). A session consisted of (1) finding a rough estimate of their threshold at each frequency, followed by (2) a systematic assessment of their threshold at each frequency with a staircase procedure (described below). First, the approximate threshold of each participant was estimated by the method of limits performed with a verbal yes/no task. This initial threshold estimate was used to guide the settings for the upper and lower bounds of the staircases. Next, the detailed threshold measurement began with a block of practice trials at 1000 Hz, followed by 23 experimental blocks (one for each frequency). At the beginning of each frequency block, participants heard a sample tone, which was the same intensity as the loudest sound presented in the given staircase. If there were any frequency blocks where the staircase procedure failed to collect conclusive thresholds (described below), participants returned to re-do the necessary frequency blocks. During the plugged and masked listening conditions, participants were subjected to the ear plugs or masking noise soon after entering the room, and completed the rest of the session in that state.

The threshold measurement task was two-interval forced choice: participants were asked to indicate in which of two time intervals the sound had played. Five hundred msec after initiating a trial with a press of the space bar, a "1" appeared on the display for one second, followed by an interval of 500 msec with a blank screen, and then a "2" appeared on the display for one second. One of the sound files was always played in either the first or second interval, and no sound was delivered in the other interval. Participants could then indicate which interval they heard the sound in by responding with a "1" or "2" on the numeric pad of the keyboard. Participants then received feedback before cycling back to the trigger screen.

Threshold was assessed for each frequency with two randomly-interleaved staircases, one ascending and one descending. Each staircase was an up-down transformed staircase (UDTR), following a 1-up, 2-down rule so that it converged on the 70.7% threshold [12]. This method of threshold assessment allows a high amount of efficiency, while maintaining sufficient complexity to avoid participants anticipating stimulus values. The step size was 3 dB of attenuation, and 10 steps could occur before the attenuation would no longer increase or decrease. Each block ended when 7 reversals occurred for both the ascending and descending staircases, or when responses continually drove the stimulus values to the upper or lower bounds of the staircases.

Despite pre-testing the thresholds the procedure did not converge to a clear threshold in every single frequency block. In the cases where it did not converge, participants came back and collected data with a different set of upper and lower bounds. Most of the time these re-collections were successful. When an exact threshold was still not determined, the softest sound in the staircase was used as the threshold if the stimuli were still too loud, and the loudest sound in the staircase was used if the stimuli were still too soft.

3. RESULTS

Figure 2 shows that the lowest threshold (i.e., maximum sensitivity) in both the open and plugged conditions occurred at 1170 Hz. In the masked listening condition, the lowest threshold occurred at 1370 Hz, though the threshold value at 1170 Hz was very similar. At frequencies below or above these frequency values, greater intensity is required to detect the sound output by the bonephones. At the lower frequencies, the intensity required for detection increases basically monotonically. At the higher frequencies, the trend is much more variable. The masked listening curve is similar to the open listening condition curve, but shifted upwards. The plugged curve is similar to the open curve but shifted down for the frequencies up to 2500 Hz, where the plugged curve crosses the open and follows a slightly different pattern.

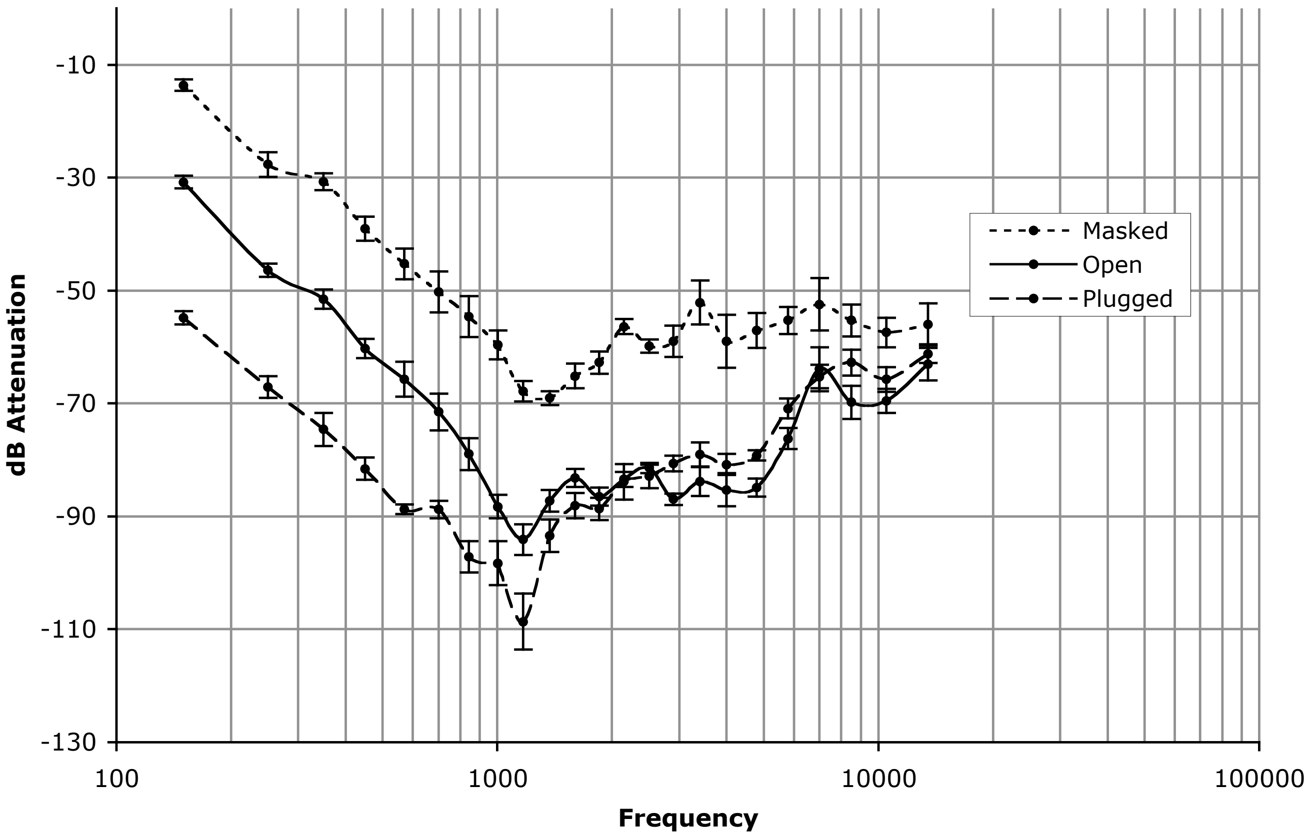

Figure 2. *Threshold of audibility with bone conduction headphones under conditions of Masked, Open, and Plugged ears.*

4. DISCUSSION

Relative to previous research, the threshold of audibility had a similar overall shape, for both the plugged and open condition, despite differences in procedure and apparatus. Similarities include the fact that the lowest threshold (most sensitivity) occurred within the frequency range of speech, and that plugging the ears lowered the threshold up until about 2000 Hz, where the open and plugged curves converged. The high number of frequencies tested gives a more detailed description of the threshold curve than previous research, which seems to be especially important in the frequencies above 1370 Hz. Although the lowest thresholds are in the frequency range of speech, the lowest thresholds within that range occur at lower frequencies than indicated by previous research. This may be due to characteristics of the bonephones' frequency response, which is discussed in greater detail below.

This is the first investigation that the authors are aware of that has considered the variable effect of binaural masking on bone conduction thresholds at different frequencies. One might expect the threshold curves in different listening conditions (i.e., masked, open, or plugged ears) to be identical in shape, but simply shifted upward by roughly the same amount at all frequencies. Although the curve shapes are very similar (more so in the lower frequencies), and the masked curve is certainly shifted up at all frequencies, masking seems to have less of an effect at the higher frequencies (7000 Hz and above). Although this may be attributable to a difference in perceptual response, an alternative explanation for this is the reduced power at higher frequencies in pink noise.

It is important to note that by measuring attenuation at the level of the sound file, the stimulus is specified at the level of input into the bonephones, before the frequency response of the device is accounted for. So, the data presented here do not represent the threshold of bone-conduction hearing, per se. Rather, these thresholds are an applied assessment of how much intensity needs to be driven into these particular bonephones in order for a listener to hear a sound, at a variety of frequencies in various practically relevant listening conditions. That is, the present data provide an overall system response curve, including the bonephone hardware and the listener in the system. This is perhaps more relevant to display designers, rather than to psychologists more interested in how the auditory system works.

Due to the nature of our stimulus specification, it may be useful to think of the threshold as an equalization curve for sounds to be perceived of equal loudness. In addition to adjustments that need to be made across frequencies when the ears are open in a quiet environment, the curve specifies how equalization should change when the ears are plugged to protect hearing, and when the environment has ambient noise.

These equalization specifications are useful because they can be used to optimize audio for the bonephones under the various listening conditions. The variability across frequencies and listening conditions in the data is evidence that this optimization is needed; if the curves were flat across frequencies and did not depend on listening conditions, then equalization would not be beneficial. These curves suggest that frequencies between 570 and 1850 Hz should be attenuated, while frequencies at the low and high ends should be boosted. Indeed, the subjective listening experience with the bonephones is that the there is not enough low-frequency sound, and too much in the midrange frequencies. The need for this optimization is not surprising, given the large differences between air and bone-conduction, both in terms of the physical properties of the device and the auditory mechanism through which sound travels.

This equalization shows what is necessary for flatter, less colored sound to be delivered to the listener wearing the bonephones. What is perceived to be a flat equalization curve must be understood before any purposeful spectral changes are made to the sound for whichever effects may be desired. This equalization is the first in a series of investigations that can eventually lead to a description of the signal processing filters that need to be applied for the spatialization of sounds played through the bonephones. Although investigations of spatial audio using bone-conduction are rare due to a common belief that crosstalk is too great to achieve effective binaural separation, the present authors and colleagues have recently shown that there is at least some spatial separation that can occur with these bonephones, and thus that spatialization of sounds through bonephones is a worthwhile goal to pursue. In particular, we found evidence for binaural separation evidenced by improvements in performance on an applied spatial audio task as a function of interaural differences [13]. In addition, bonephones have successfully delivered spatialized navigational cues in an audio wayfinding system [14]. Future research that needs to be done includes more immediate steps between basic equalization and complete implementation of signal processing for spatial audio, such as thresholds of acoustic cues leading to lateralization. Future research may also benefit from extending the range of tested frequencies up into the range that Corso [8] investigated. These ultra-hearing frequencies could be used to extend the range of sonifications with pitch mappings [15].

5. REFERENCES

[1] G. von Bekesy, *Experiments In Hearing*. New York: McGraw-Hill Book Company, Inc., 1960.

[2] J. Tonndorff, "Bone Conduction," in *Foundations of Modern Auditory Theory*, J. V. T. (Ed.), Ed. New York: Academic Press, 1972, pp. 195-237.

[3] B. N. Walker and J. Lindsay, "Navigation performance with a virtual auditory display: Effects of beacon sound, capture radius, and practice," *Human Factors*, in press.

[4] S. S. Stevens and H. Davis, *Hearing: Its Psychology and Physiology*. New York: John Wiley & Sons, Inc., 1938.

[5] G. A. Studebaker, "Placement of vibrator in bone-conduction testing," *Journal of Speech & Hearing Research*, vol. 5, pp. 321-331, 1962.

[6] "ISO 389-3 - Acoustics - Reference zero for the calibration of audiometric equipment - Part 3: Reference equiv alent threshold force values for pure tones and bone vibrators," International Organization for Standardization, Geneva, Switzerland 1994.

[7] D. W. Robinson and M. S. Shipton, "A standard determination of paired air- and bone-conduction thresholds under different masking noise conditions.," *Audiology*, vol. 21, pp. 61-82, 1982.

[8] J. F. Corso and M. Levine, "Pitch discrimination at high frequencies by air- and bone- conduction," *American Journal of Psychology*, vol. 78, pp. 557-566, 1965.

[9] N. A. Watson, "Limits of audition for bone conduction," *Journal of the Acoustical Society of America*, vol. 9, pp. 294-300, 1938.

[10] J. D. Hood, "Bone conduction: A review of the present position with especial reference to the contributions of Dr. Georg von Bekesy," *Journal of the Acoustical Society of America*, vol. 34, pp. 1325-1332, 1962.

[11] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and models*, Second ed. Berlin: Springer, 1999.

[12] H. Levitt, "Transformed Up-Down Methods in Psychoacoustics," *Journal of the Acoustical Society of America*, vol. 49, pp. 467-477, 1971.

[13] B. N. Walker, R. M. Stanley, N. Iyer, B. D. Simpson, and D. S. Brungart, "Evaluation of Bone-Conduction Headsets for Use in Multitalker Communication Environments," *Submitted for publication*, 2005.

[14] B. N. Walker and J. Lindsay, "Navigation performance in a virtual environment with bonephones," *Manuscript submitted for publication*, 2005.

[15] B. N. Walker, "Magnitude estimation of conceptual data dimensions for use in sonification," *Journal of Experimental Psychology: Applied*, vol. 8, pp. 211-221, 2002.