COMMENTARY

Absolute and relative pitch processing in tone learning tasks
Sandra E. Trehub

Department of Psychology, University of Toronto, Canada

Saffran offers her results as support for ‘the hypothesis that absolute pitch is present in infancy’ (Abstract). Before evaluating her claim, it is reasonable to review some facts about absolute pitch (AP) and about AP possessors. Since the phenomenon was first submitted to scientific scrutiny (Stumpf, 1883), AP has been defined as the ability to identify a specific tone by its musical name (e.g. C) or to produce it (by singing or adjusting a tone generator) without using a reference pitch. The assumption is that individuals with AP have a fixed, or stable, internal template of note names (e.g. C, D, E, F, G, A, B) attached to specific frequencies or pitches (Ward, 1999). When AP possessors wake up in the morning, they can name any isolated tone whose pitch corresponds to a note on the musical scale. Indeed, such unconventional test procedures have been used to preclude adults’ reliance on reference tones (Petran, 1932). More commonly, the test context is configured to minimize the availability of reference tones (Costall, 1932). More commonly, the test context is configured to minimize the availability of reference tones (Costall, 1985).

AP is considered special because of its relative rarity (1 in 10,000: Takeuchi & Hulse, 1993). In general, trained musicians or musically educated amateurs have relative pitch (RP) rather than AP. Instead of the stable template associated with AP, RP involves a movable template corresponding to precise relations among notes of the scale (Ward, 1999). RP possessors recognize the musical interval (pitch distance) between any pair of tones (e.g. perfect fifth, major third), which enables them to name the second of two tones when the first is named. When they hear a single (unnamed) tone, they can produce a second tone to yield a specified interval. The less developed RP skills of musically untrained individuals can be evaluated in other ways: judging a well-known melody as in or out of tune (Drayna, Manichaikul, de Lange, Snieder & Spector, 2001), or judging whether the interval between two tones matches an interval in long-term memory (e.g. the first two tones of Here Comes the Bride; Smith, Kemler Nelson, Grohskopf & Appleton, 1994).

Tasks that involve musical labels for tones or intervals are necessarily limited to musically sophisticated listeners who have well-established associations between verbal labels (note or interval names) and pitches (for AP) or pitch combinations (for RP). Tasks without such labelling requirements have provided information about untrained adults’ long-term memory for specific pitches (Levitin, 1994; Schellenberg & Trehub, in press) or intervals (Drayna et al., 2001). In any case, concepts related to AP and RP are largely irrelevant to Saffran’s current study of statistical learning or to her previous explorations of tone learning (Saffran & Griesentrog, 2001; Saffran, Johnson, Aslin & Newport, 1999), which did not explore long-term memory. With no basis for attributing AP or RP to the infants or adults in these studies, how can we interpret the findings? We can ask, instead, whether infants and adults used absolute or relative cues to solve the task. Can we accept Saffran’s contention that infants capitalized on absolute pitch cues and that adults capitalized on relative pitch cues?

Musicians in Saffran and Griesentrog (2001) should have obtained reasonable performance levels on the basis of their RP abilities. Instead, they achieved 63% correct on the RP task and 56% on the AP task (non-musicians obtaining 57% and 52%), which raises concerns about Saffran’s task designations. Non-musicians in Saffran et al. (1999) achieved roughly 65% correct on a comparable AP task with different tone sequences, which indicates that the component tones of training and test sequences can affect performance. It is likely that small pitch distances between the component tones of ‘words’ in Saffran et al. (1999) contributed to the coherence and memorability of those sequences, in line with perceptual grouping principles (Bregman, 1990;
Deutsch, 1999). Regardless, Saffran’s conclusion that adoption of a relative pitch strategy is at odds with their failure to perform better on RP than on AP tasks in the present study. Her use of multiple test trials may have impaired adults’ performance, perhaps differentially, on AP and RP tasks. Indeed, the presentation of intervening tones is often used to ‘erase’ pitch memory in AP tasks (Ward, 1999).

If infants detected differences between words and parts-word on the basis of absolute pitch cues, they should have succeeded on RP as well as AP tasks. The test items in both tasks included sequences that were presented frequently during the familiarization period. The principal difference between AP and RP tasks was the non-distinctiveness of relative pitch cues among test items in the AP task. Specifically, the AP task had distinctive absolute pitch cues, and the RP task had distinctive absolute and relative pitch cues, along with distinctive contour cues. These factors make it impossible to pinpoint the basis for infants’ or adults’ performance on the AP and RP tasks.

Saffran’s proposed developmental shift from absolute to relative pitch processing has parallels in infants’ shift from language-general to language-specific processing of speech sounds (Werker & Tees, 1999). Her proposal is also consistent with the hypothesized critical period for the acquisition of AP (Takeuchi & Hulse, 1993). However, unlike very early exposure to a second language, which guarantees accent-free speech (Flege & Fletcher, 1992), early musical training (before 6 years) is necessary but not sufficient for AP (Baharloo et al., 1998). Genetic factors also make important contributions (Baharloo, Johnston, Service, Gitschier & Freimer, 1998; Baharloo, Service, Risch, Gitschier & Freimer, 2000), as they do for RP (Drayna et al., 2001).

In short, Saffran’s proposed developmental shift requires much more evidence than that presented. For example, it remains to be determined whether there are developmental changes in pitch memory and in the relative weighting of absolute and relative pitch cues. Despite the prevailing view that adults typically focus on distinctive absolute and relative pitch cues, there is increasing evidence that they encode novel melodies in absolute rather than relative terms (Dowling, 1999). Finally, it is possible that the critical period for AP has less to do with age-related changes in pitch memory than with developmental changes in the ease of acquiring arbitrary associations between verbal labels and pitches.

References


RESPONSE

Birds do it – why not babies?

Jenny R. Saffran

Waisman Center and Department of Psychology, University of Wisconsin-Madison, USA

Like many species of songbirds and non-human primates, we suggest in the current paper and elsewhere (Saffran & Griepentrog, 2001) that human infants have the capacity to represent the absolute pitches of sounds, which we defined as ‘the encoding of a pitch independent of its relation to other sounds’. Whether or not this mode of perceiving the auditory world is related to the rare ability to label pitches remains unknown, and studying the ontogeny of pitch labeling is not the intent of this line of research. Instead, our interest lies in characterizing the cues utilized by learners at different points in development. That is, to what extent do infant listeners share the perceptual primitives – absolute pitches – observed frequently in non-human species, and how does this capacity change over age and experience?

We suggest that our results support the claim that infants preferentially represent absolute pitches whereas adults preferentially represent relative pitches, given the task of segmenting a continuous stream of tones. Directly pitting these cues against one another is difficult, given that any alteration in absolute pitches necessarily alters the relative pitches. It is for this reason that we chose the tact of paired experiments. Each experiment was designed such that success on the test would require the use of either one type of pitch cue or the other. As in any experiment, we can only interpret successful discrimination; in the case of the Saffran and Griepentrog (2001) results, infants successfully discriminated when AP cues were required, and adults successfully discriminated when RP cues were required. Because this is a different pattern of successes as a function of development, we interpreted these data as indicating a shift in perceptual prioritization.

Importantly, as noted in our papers, we do not claim that these cues are used to the exclusion of other types of information. Indeed, adults in the current study do show some success at tracking sequences of absolute pitches, consistent with other literature demonstrating the use of AP cues by adult non-musicians (Levitin, 1994; Halpern, 1989). Moreover, infants can apparently be induced to track RP cues in our task if AP cues are rendered non-predictive; preliminary data from our lab indicate that when RP cues are consistent but AP cues are highly variable, 8-month-old infants can track relative pitch sequences in our stimuli (Saffran, Reec & Niebuhr, in progress). This pattern of results suggests a complicated interplay between the perceptual capacities of learners and the cues afforded by auditory stimuli: adult listeners are more successful at tracking absolute pitches as the materials become easier to represent, while infant listeners are more successful at tracking relative pitches when absolute pitch cues are rendered less informative.

We agree with Sandra Trehub that the component tones of the training and test sequences can affect performance; this is clearly evident in the difference in performance between the current study and Saffran and Griepentrog (2001). However, it is not the case that non-musicians in another of our studies (Saffran, Johnson, Aslin & Newport, 1999) achieved roughly 65% on a comparable AP task, as suggested by Trehub. That task was not designed to isolate AP cues; RP cues were also available for some of the test discriminations used by Saffran et al. (1999). Comparing performance levels from Saffran et al. (1999) with the results of our more recent studies, in which no RP cues were available for test discrimination, does not provide a valid indicator of the use of different types of pitch cues. It is also important to note that infants’ failure in the RP tasks is unlikely to be due to inherent difficulties with the coherence and memorability of the ‘words’, given that adults can successfully learn these materials (while failing to learn others on which the infants succeed).

Is it the case that infants should have succeeded on the RP task given access to AP cues? We suggest not. While words and part-words did necessarily differ in their AP cues, the design ensured that the transitional probabilities between absolute pitches in the part-words and the
words are identical, limiting the usefulness of AP as a cue for discrimination. Infants’ failure to discriminate suggests that the equating of absolute pitch familiarity across test items provided a sufficient control. Relative pitch cues – which provided a strong cue for discriminating words from part-words – also failed to lead to discrimination, which we interpret as a lack of evidence for the use of RP cues given these materials as input.

We thus conclude that infants do have access to absolute pitch information. At the same time, we absolutely agree with Trehub that much more remains to be learned about the use of pitch cues across development. We suggest that all of our learners can use both types of pitch cues, but that the circumstances under which each type of cue is used are driven by different factors at different ages. On this view, the interesting question is what drives the relative strengths of the many cues which co-exist in the perceptual world, and why might this weighting change over development? Some avian species can be induced to switch from prioritizing absolute pitch to using relative pitch as a function of the structure of the learning materials (MacDougall-Shackleton & Hulse, 1996); we suspect that humans share this capacity to use both types of pitch cues, as well as the flexibility to switch perceptual primitives as the situation demands.

References