AUGUST 19 2025

A re-examination of duplex perception with musical chords [a)]

Brian Roberts

J. Acoust. Soc. Am. 158, 1246–1253 (2025)
https://doi.org/10.1121/10.0038976

A re-examination of duplex perception with musical chords[a)]

Brian Roberts[b)]
School of Psychology, Aston University, Birmingham B4 7ET, United Kingdom

ABSTRACT:
In duplex perception, an acoustic element differing from the others in receiving ear or form (e.g., harmonic complex or sinusoidal) contributes simultaneously to two distinct percepts. Speech has received the most attention, but duplex perception also occurs with musical chords. By one account, a specialist processing "module" (e.g., phonetic) has priority access to acoustic information (precedence) that is not subject to scene-analysis constraints. Precedence received initial support from claims of better performance for speech than non-speech judgments of the same stimuli, but experiments controlled for criterion differences and unintended cues challenged this interpretation. This approach is extended here to three-note chords comprising tonic and fifth complexes and a sinusoidal third (30-dB presentation range) defining the mode (major/minor). In experiment 1, listeners first heard a chord and identified its mode; chords were then preceded by two successive sinusoids—one matching the third, the other mistuned—and listeners identified the matching tone. In experiment 2, for tone discrimination, the tonic and fifth were replaced by a single complex crafted to produce equivalent masking of the third but to remove an unintended mode cue. The notion of precedence was not supported; there was no evidence of better performance for musical than non-musical judgments of the same stimuli.
 https://doi.org/10.1121/10.0038976

(Received 14 March 2025; revised 23 July 2025; accepted 26 July 2025; published online 19 August 2025)

[Editor: Emily Buss] Pages: 1246–1253

I. INTRODUCTION

Exclusive allocation is considered a core principle of scene analysis in vision and hearing, according to which a given sensory element can only be allocated perceptually to one object or stream at the same time. However, in an apparent violation of this principle, there are occasional circumstances in which a given element can contribute to two distinct percepts at once, a phenomenon known as duplex perception. The duplex perception of speech (Rand, 1974; Liberman, 1982) has received the most attention, but duplex perception has also been found for musical chords (Pastore *et al.*, 1983; Collins, 1985) and even for environmental sounds, such as the sound of a closing door (Fowler and Rosenblum, 1990). Duplex perception can also occur for a mistuned harmonic in a complex tone: The mistuned component may be heard as a separate tone but may nonetheless still contribute to the global pitch of the complex tone (Moore *et al.*, 1985, 1986). Demonstrations of duplex perception have been taken to indicate the simultaneous operation of distinct perceptual systems and thus potentially to provide insights into the architecture of perceptual processing. The aim of the experiments reported here was to explore further the duplex perception of chords by comparing performance on tasks involving musical and non-musical judgments of these stimuli while better controlling for criterion effects and unintended cues.

The duplex perception of speech was originally demonstrated using elements of a simplified synthetic consonant-vowel (CV) syllable that, when presented to listeners dichotically, gave rise to the simultaneous perception of a speech sound and a non-speech sound (Rand, 1974; Liberman, 1982). Experiments of this kind, discussed in more detail below, were used by researchers at Haskins Laboratories to argue that speech perception involves the operation of a specialized phonetic module pre-emptive in its use of sensory information—taking what it needs first without being subject to the usual scene-analysis constraints (Bregman, 1990)—and so takes precedence over non-speech perception (Whalen and Liberman, 1987; Liberman and Mattingly, 1989). The argument is based primarily on claims of differences in responses to the speech and non-speech components of a duplex percept (Whalen and Liberman, 1987), but this result is contentious for procedural reasons, and others have reported little or no difference in performance between corresponding speech and non-speech tasks (Bailey and Herrmann, 1993). An analogous difference in performance between the musical and non-musical components of a duplex percept has been reported for chords (Hall and Pastore, 1992), but this claim has not yet been critically re-examined.

Most studies of duplex perception have used a dichotic configuration in which one ear receives the non-distinguishing parts constituting most of the stimulus (the "base"), but the critical element carrying the information

---

[a)]The experiments reported here correspond to reconsidered versions of experiments 4 and 5 in the unpublished doctoral thesis of Bruce (2000). A poster presentation on this research was given at the Auditory Science Meeting (Cambridge, UK, September 2024).
[b)]Email: b.roberts@aston.ac.uk

https://doi.org/10.1121/10.0038976

needed to identify the whole stimulus is presented contralaterally, in the correct temporal alignment. Under these conditions, the critical element perceptually completes the base (i.e., the syllable or chord can be identified in the ear receiving the base) but is heard simultaneously as an additional (non-speech or non-music) sound in the other ear. In the classic demonstration of the duplex perception of speech (Rand, 1974; Liberman, 1982), synthetic CV syllables were created for which the only cue disambiguating the identity of the voiced stop (/b/, /d/, or /g/) was the initial third-formant (F3) transition, which rose or fell to varying degrees. That element not only supported the consonant percept but was also heard simultaneously as a non-speech "chirp." Similarly, presenting the base of a chord (perfect fifth) to one ear and the distinguishing third to the other led to the perception of a complete chord in the ear receiving the base (i.e., the chord was heard as major or minor) and the third was heard as an additional, isolated note in the other ear (Pastore *et al.*, 1983; Collins, 1985). Note that the two modes of listening involved the following: Identifying the syllable or chord involved synthetic listening, whereas discriminating the distinguishing chirp or third involved analytic listening.

Whalen and Liberman (1987) created a variant of the test paradigm for exploring the duplex perception of speech that did not involve dichotic presentation or any other spatial cues. This was done by introducing a difference in source properties between the syllable base (buzz-excited formants) and the distinguishing element (sine-wave analogue of the F3 transition). Using this approach, they explored the effects of changing the intensity of the sinusoidal tone glide on the speech and non-speech percepts. The duplexity threshold was defined as the glide intensity, relative to the steady F3, above which the syllable was heard as accompanied by a non-speech sound (heard as a rising or falling whistle). Strikingly, Whalen and Liberman (1987) reported that there was a range of intensities below that required for detection of the glides (i.e., below the duplexity threshold) for which listeners were still able to distinguish the initial stop consonant as /d/ or /g/. Estimates of the difference between the duplexity threshold and the threshold for syllable discrimination ranged from at least -4 dB (experimental data) to around -20 dB (anecdotal reports by those authors based on their own listening). Whalen and Liberman (1987) regarded this difference in thresholds as evidence for the precedence of the speech mode in perception. Specifically, speech perception took first what was required from the sensory evidence to support the consonant judgment, leaving only a "residual" for the non-speech judgment (rising vs falling glide). Hence, the non-speech judgment fails first with increasing attenuation of the glide. If both systems instead either had full access to the available evidence or shared it equally, there should be no difference in thresholds.

On two grounds, Bailey and Herrmann (1993) questioned the claim by Whalen and Liberman (1987) that there was a range of subthreshold levels [so-called subliminal perception; see Hall and Pastore (1992)] over which the glide nonetheless influenced consonant identity in the base+glide syllables. First, Whalen and Liberman (1987) used the method of adjustment to measure the duplexity threshold, an approach for which the threshold depends on the criterion that the listener sets for detection. Further, the glide intensity on those trials always began below the threshold, which is known to lead to an overestimation of threshold [see, e.g., Bregman (1990), p. 55]. Hence the criterion was likely more conservative for discriminating glide direction than syllable identity. Second, Bailey and Herrmann (1993) argued that properties of the isolated tone glides (rising vs falling) might themselves provide unintended information about consonant identity, owing to associative learning. They used a two-alternative forced-choice method, which is less susceptible to criterion effects, and found that their estimates of the duplexity threshold—and consequently of the difference between the thresholds for duplexity and syllable identification—were much lower than those reported by Whalen and Liberman (1987). Further, they demonstrated that listeners could reliably label rising and falling tone glides as /g/ and /d/, respectively, indicating an unintended cue to consonant identity that did not require integration of the syllable base and glide. Overall, they concluded that there were no glide levels that did not evoke the percept of a chirp-like sound but which nonetheless allowed the consonant to be identified.

Hall and Pastore (1992) created an analogous non-spatial variant of duplex perception for musical chords, using a chord base (perfect fifth) composed of two harmonic complex tones (synthesized piano sounds) and a sine-wave analogue of the distinguishing third (same attack and decay envelope as for the complex tones). Like Whalen and Liberman (1987), they reported a considerable difference between the thresholds for duplexity and chord identification as the level of the third was varied, ranging from an average of -9 dB (experiment 1) to -18 dB (experiment 2). Hall and Pastore (1992) attributed the superior performance on chord identification to the effects of the highly learned schemas of Western music, rather than to the existence of a specialized module for music perception. In other words, modes of processing based on "figurally good" objects, such as complete chords or syllables, take precedence over other kinds of processing. In their first experiment, a yes/no task was used to measure tone detection, but a same/different task was used to measure chord discrimination. Therefore, the differences in threshold may have arisen, at least in part, from differences in detection criterion between the two tasks. Their second experiment replaced chord discrimination with chord identification and used a two-interval forced-choice method to measure tone detection, for which little or no criterion effect would be expected. However, both of their experiments used only one instance of the perfect fifth (no variation in fundamental frequency, F0) as the chord base, and so there was ample opportunity for listeners to learn the association between the frequency of the sinusoidal third and the mode of the chord.

https://doi.org/10.1121/10.0038976

Extending the approach of Bailey and Herrmann (1993), the two experiments reported here provide a re-examination of the duplex perception of musical chords in a context where criterion differences between tasks and opportunities for associative learning of unwanted cues were minimized. Using more reliable procedures to measure thresholds greatly reduced differences in performance between chord identification and tone discrimination. The notion of precedence as the mechanism underpinning duplex perception was not supported.

II. EXPERIMENT 1

Listeners made musical (mode-identification) and tone-discrimination judgments using the same stimuli. Two-interval forced-choice methods were used for both tasks to limit criterion effects and the chord stimuli were drawn from a set spanning a range of F0 frequencies to reduce the likelihood of associating a given third with a particular mode.

A. Method

1. Listeners

Ten listeners took part; all were adult volunteers (age range: 18–30 years), who reported normal hearing and were paid for participating. Most listeners were Psychology students at the University of Birmingham; the others were local residents involved in the music scene. All listeners either played or had played a musical instrument and had obtained appropriate qualifications (e.g., Royal College of Music grades). A screening test was given before the main experiment to confirm that all listeners could reliably identify major and minor chords. The screening stimuli differed from those used in the reference condition of the main experiment only in that the third was a complex tone (like the tonic and fifth) and the chords were drawn from a narrower F0 range (tonic F0: 400–424 Hz). To assist labelling, the screening began with a brief demonstration of exemplar major and minor chords, and listeners were reminded of the typical cultural associations of these modes in Western music with happiness and sadness, respectively [e.g., Burns and Ward (1982)]. All listeners exceeded the criterion for inclusion of 70% accuracy for mode identification. The ability of listeners to perform the tone-discrimination task (described below) was not screened.

2. Stimuli and conditions

The same chord stimuli were used for the musical and tone-discrimination judgments, but the stimulus configurations differed. All stimuli were presented diotically. In the mode-identification task, a single chord was presented on each trial. The tonic and fifth were complex tones comprising six harmonics added in sine phase and with a spectral roll-off of 6 dB/octave. The note representing the third was a pure tone. As for Hall and Pastore (1992), this arrangement was analogous to that of Whalen and Liberman (1987), for which the base comprised buzz-excited formants (harmonic complex) and the distinguishing F3 transition was a pure-tone glide. These triads closely approximated equal temperament, differing only in that the ratio of F0 frequencies for the tonic and fifth was set exactly at 2:3 [see Vos (1982)]. Although the frequencies used did not correspond to specific notes on any scale used in Western music, the relative frequencies were appropriate for the perception of major and minor chords. Specifically, the frequencies of each major and minor third were set to have ratios, relative to the tonic F0, of 1:1.26 and 1:1.19, respectively.

To prevent listeners from identifying mode based on the frequency of the third, a set of 11 thirds was used (400 Hz–600 Hz, in 20-Hz steps) and chords were constructed such that each third was equally likely to correspond to the major or minor mode. The token used on any given trial was selected randomly, and with equal likelihood, from the available set. The wide range of thirds also ensured that the F0 ranges for the tonic and fifth showed considerable overlap across the two modes (tonic F0, 317.5 Hz–476.2 Hz [major] and 336.1 Hz–504.2 Hz [minor]; fifth F0, 476.2 Hz–714.3 Hz [major] and 504.4 Hz–756.3 Hz [minor]). Each chord had a trapezoidal amplitude envelope and was 500 ms long, including 50-ms linear onset and offset ramps. In the reference condition, the pure-tone third was presented at the same level as the F0 components of the tonic and fifth—75 dB sound pressure level (SPL)—giving an overall presentation level for the chord of 80 dB SPL. In the other conditions, the attenuation of the third ranged from 5 to 30 dB in 5-dB steps. Hall and Pastore (1992) used the same range of attenuation for their pure-tone thirds.

In the tone-discrimination task, each stimulus was a sequence of three sounds—two pure tones and a chord. The pure tones had the same duration and amplitude ramps as their counterpart within the chord; all sounds were separated by 0.5-s silences, giving a total stimulus duration of 2.5 s. One of the pure tones (the target) was identical in frequency to the third of the chord, and the other (the foil) was mistuned by $\pm 5\%$. This extent of mistuning was chosen to be broadly comparable with the frequency separation for a major third and minor third with respect to a common perfect fifth; this was intended to help optimize the comparison between the synthetic and analytic modes of listening required for the two tasks. The level of the pure tones was set equal to that of the third in the chord, as appropriate for the level of attenuation used on any given trial. The pure tones remained clearly audible at maximum attenuation (30 dB, giving 45 dB SPL).

All stimuli were generated using MITSYN (Henke, 1997) at a sampling rate of 16 kHz, stored on disk, and played back via a 16-bit digital-to-analogue converter. They were low-pass filtered at 5.2 kHz (roll-off = 100 dB/octave) and presented through Sennheiser (Hannover, Germany) HD 480-13II earphones. The output level was set using a programmable attenuator (0.25-dB steps) and calibrated using a sound-level meter [Brüel and Kjaer (Naerum, Denmark) type 2209] coupled to the earphones via an artificial ear (Brüel and Kjaer, type 4153).

3. Procedure

Stimuli were played to the listeners in a single-walled sound-attenuating booth [Industrial Acoustics (Winchester, UK) 401A] housed within a quiet room. The two tasks were run in a single session but separately, to allow listeners to optimize their listening strategy for the task in hand. The mode-identification task was always completed first. There was a short rest break between tasks, and each began with a warm-up comprising one block of the appropriate stimuli, the data from which were discarded. In total there were 560 trials for the main experiment, with 280 per task (2 chords $\times$ 7 attenuations $\times$ 20 blocks); stimulus order was randomized anew for each block. For the tone-discrimination task, the 20 blocks of trials comprised five repetitions of the four permutations of note order (target, +5% mistuned foil; target, –5% mistuned foil; +5% mistuned foil, target; –5% mistuned foil, target).

In the mode-identification task, listeners indicated whether they heard the chord as major or minor by pressing "M" or "N" on the keyboard, respectively. Tone discrimination was measured using a two-alternative forced-choice task, in which listeners indicated whether the note within the chord was present in the first or second interval by pressing key "1" or "2," respectively. Listeners were encouraged to make their judgments quickly and accurately, but their responses were not time limited. On each trial, feedback on the accuracy of the response was given on-screen immediately afterwards for both tasks. The next trial began 0.5 s after it was initiated by the listener.

B. Results and discussion

Figure 1 shows the mean percentage scores (and inter-subject standard errors) for each task across conditions. The top panel shows the overall results for each task when collapsed across stimulus mode; results for mode identification and tone discrimination are displayed using filled circles and squares, respectively. The middle and bottom panels show the results separately for major and minor stimulus configurations in the mode-identification and tone-discrimination tasks, respectively; results for major and minor chords are displayed using open upward and downward triangles, respectively. Results for the two tasks were analyzed together using a three-way repeated-measures analysis of variance (ANOVA); the measure of effect size reported here is partial eta squared (η_p^2). The independent variables were task (two levels: mode identification or tone discrimination), stimulus mode (two levels: major or minor), and attenuation of the pure-tone third (seven levels: 0–30 dB). The outcomes of the ANOVA are summarized in Table I. All statistical analyses were computed using SPSS STATISTICS (version 26, 2019; IBM Corp., Armonk, NY).

Attenuation was the only factor associated with a significant main effect ($p < 0.001$). Mean performance was good in both tasks for the reference case (0 dB, $\geq$90%) but fell progressively as the level of the pure-tone third was lowered (Fig. 1, top panel). There was a highly significant interaction between task and attenuation ($p < 0.001$), corresponding to the shallower function for tone discrimination than for mode identification. In the latter case, accuracy fell to chance for the highest levels of attenuation but remained above it

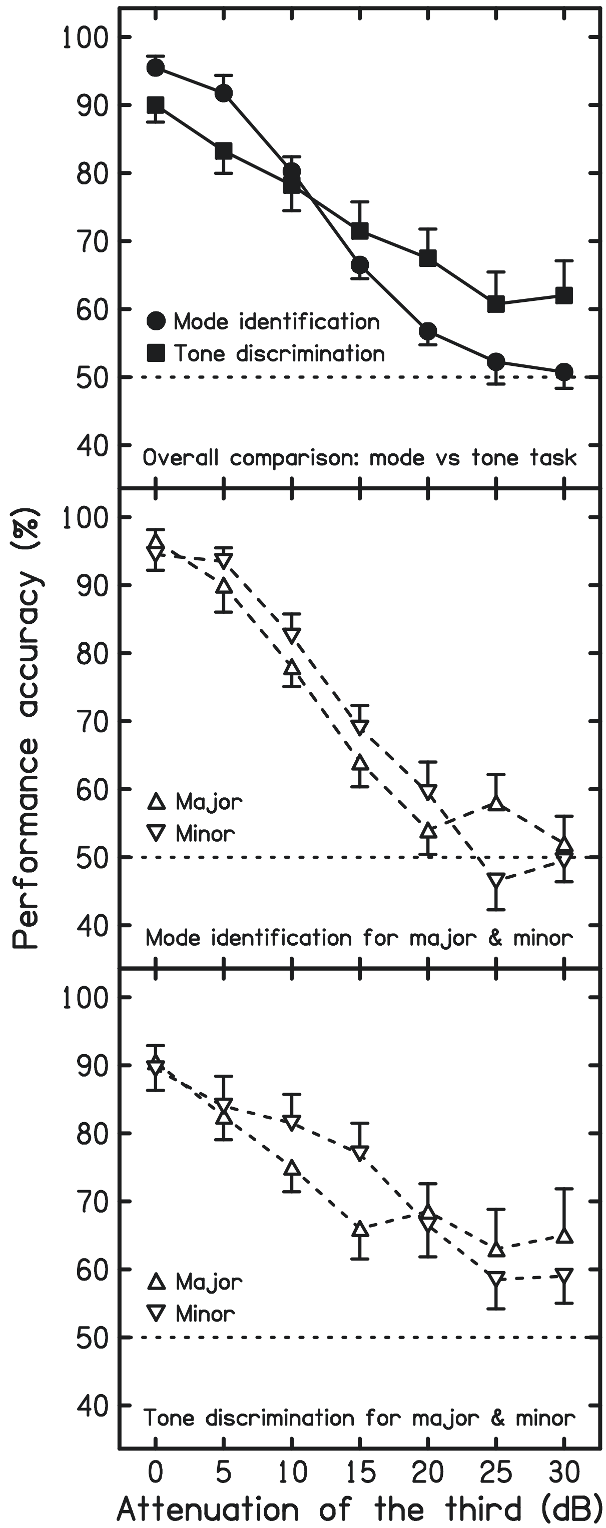

FIG. 1. Results of experiment 1. (Top panel) Results for each task, collapsed across stimulus mode. (Middle and bottom panels) Results are shown separately for major and minor stimulus configurations in the mode-identification and tone-discrimination tasks, respectively. For greater clarity, error bars are shown in one direction only. Dotted lines indicate chance performance (50%).

TABLE I. Summary of the three-way repeated-measures ANOVA for experiment 1. Task had two levels (chord identification or tone discrimination), stimulus mode had two levels (major or minor), and attenuation (intensity of the third) had seven levels (0–30 dB, in 5-dB steps). All significant terms are shown in bold.

Factor	*df*	*F*	*p*	η_p^2
Task (T)	(1, 9)	0.810	0.392	0.083
Mode (M)	(1, 9)	0.482	0.505	0.051
Attenuation (A)	**(6, 54)**	**120.988**	**<0.001**	**0.931**
T × M	(1, 9)	0.013	0.912	0.001
T × A	**(6, 54)**	**7.886**	**<0.001**	**0.467**
M × A	**(6, 54)**	**3.047**	**0.012**	**0.253**
T × M × A	(6, 54)	1.120	0.363	0.111

(∼60%) in the former. A possible explanation for this difference between tasks is discussed below. The only other significant interaction was between stimulus mode and attenuation ($p = 0.012$), which arose mainly from the tendency for minor chords to give higher scores than major chords for both tasks at mid-range attenuations (Fig. 1, middle and bottom panels). There is no obvious reason for this interaction, and it is not considered further.

Adaptive procedures are often used to estimate thresholds in two-alternative forced-choice tasks (Levitt, 1971), for which one-up, two-down and one-up, three-down staircases converge on 70.7% and 79.4% accuracy, respectively. Hence, scores in the range 70%–80% on the psychometric function are widely taken to indicate "threshold" performance. Attenuation of the third by 10–15 dB produced accuracy in this range in both tasks, corresponding at 75% performance to a difference in tone attenuation of less than 1 dB for the two tasks. This outcome is not consistent with the notion that subliminal perception took place [cf. the much larger average difference in thresholds of −9 dB for Hall and Pastore (1992), experiment 1]. Nonetheless, for the largest attenuations tested (20–30 dB), performance was better for tone discrimination than mode identification.

Note that even the most attenuated tones remained easily audible when heard in isolation, and so listeners may have been able to make a sequential comparison with the perfect fifth to ascertain which one was in the correct ratio to form a major or minor chord. Listener comments suggested that this might have been used as a cue [see also Hall and Pastore (1992)]. If so, this cue would be an imperfect one, because the frequency difference between target notes and foils (±5% mistuning) is similar to that between major and minor thirds (one semitone ≈ 6%), such that −5% on a match to the major third is close to the minor position and +5% relative to the minor third is close to the major. Hence, in principle, relying solely on this cue once the embedded third was attenuated to the point that it no longer supported chord identification could have allowed up to 100% accuracy in half the trials (foil: +5% relative to major third or −5% relative to minor third) but near-chance accuracy on the other half, setting an overall performance ceiling close to 75%. In practice, performance at the greatest attenuation used was poorer than this on average (∼60% correct), but inspection of the data for individual listeners showed that they were evenly split between those who could and could not use the strategy of sequential comparison. A revised tone-discrimination task was used in experiment 2 to address this unwanted musical cue.

III. EXPERIMENT 2

This experiment used the same stimuli as for experiment 1 in the mode-identification task, but differed in that the two complex tones making up the perfect fifth were replaced in the tone-discrimination task by a complex tone intended to provide equivalent peripheral masking of the embedded pure-tone third without providing an unwanted musical cue. Hall and Pastore (1992; experiment 2) achieved a similar goal by replacing the two complex tones making up the perfect fifth with a narrow-band noise masker. However, band-limited noise has appreciable amplitude modulation and their use of peak- rather than average-level matching to set its overall level may have contributed to the substantial difference in thresholds between tasks that they observed (–18 dB on average). Therefore, a discrete-spectrum masker was used instead.

A. Method

Ten listeners took part, six of whom also participated in experiment 1. Except where indicated, the method was the same as for experiment 1. The range of attenuations applied to the third was changed from 0–30 dB to 5–35 dB, again using 5-dB steps. This change was intended further to optimize the test range with respect to near-chance performance. For the mode-identification task, the stimuli and procedure were otherwise identical to those for experiment 1. For the tone-discrimination task, the pair of complex tones making up the perfect fifth was replaced by a ten-component "masking complex" with a flat spectrum and an F0 frequency equal to half that of the tonic (see Fig. 2). The frequencies of harmonics 2 and 3 of this complex were equal to the F0s of the original tonic and fifth, and each harmonic had the same level as the F0 components of those notes, providing a "frame" corresponding to the perfect fifth within which the pure-tone third could be added in the appropriate ratio to provide analogues of the major and minor chords used in the mode-identification task.

The masking complex had the same amplitude envelope as the pair of complexes it replaced. It was intended greatly to reduce or eliminate the opportunity for listeners to use modal information for their discrimination of the pure-tone third. Note that the masking complex constituted a perceptually fused harmonic series with an F0 one octave below that of the tonic it replaced. Consequently, the third did not correspond to a harmonic of F0, whether it was in the major or minor position with respect to components 2 and 3. Mistuning is known to promote segregation and hence was expected to increase the salience of the pure-tone third (Moore *et al.*, 1986). Although the masking complex was a

https://doi.org/10.1121/10.0038976

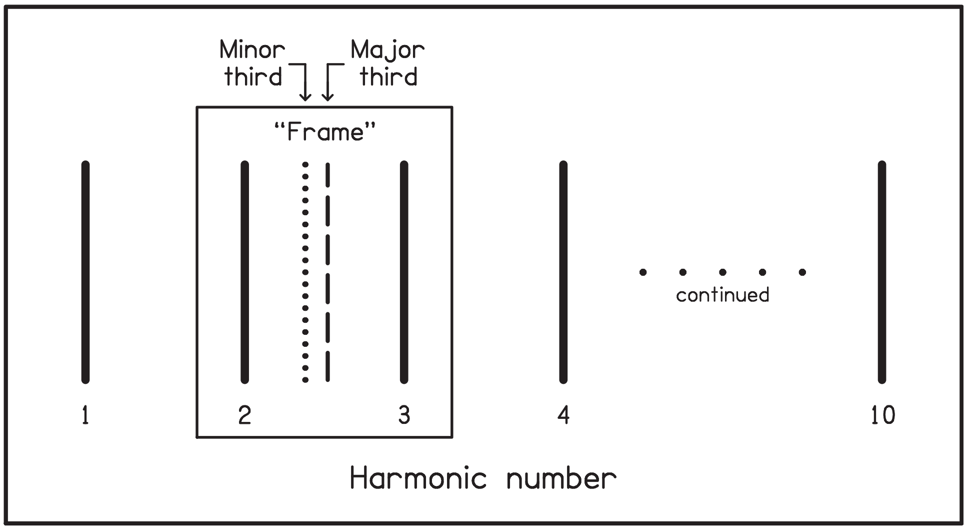

FIG. 2. Schematic depicting the stimulus configuration used in experiment 2. The "masking complex" consisted of harmonics 1–10 (solid black lines) of an F0 frequency one octave below that of the original tonic. Harmonics 2 and 3 acted as a "frame," defining the perfect fifth, within which the pure-tone third was added in either the major (dashed line) or minor (dotted line) position.

little more intense overall than the pair of complex tones it replaced, only components 2 and 3 of this complex would have had any appreciable masking effect on the third (Glasberg and Moore, 1990). This is very similar to the masking effect on the third of the original pair of complexes making up the perfect fifth, which arose primarily from the F0 components of the tonic and the fifth.

B. Results and discussion

Figure 3 shows the mean percentage scores (and intersubject standard errors) for each task across conditions. The top panel shows the overall results for each task when collapsed across stimulus mode. The middle and bottom panels show the results separately for major and minor stimulus configurations in the mode-identification and tone-discrimination tasks, respectively. The ANOVA was constructed in the same way as for experiment 1, but here the range of attenuation applied to the third was 5–35 dB. The outcomes are summarized in Table II.

Once again, attenuation was the only factor associated with a significant main effect ($p < 0.001$); mean performance was good in both tasks in the reference case (≥90%) but fell progressively until it approached chance as the level of the pure-tone third was lowered (Fig. 3, top panel). However, the pattern of interactions observed here was different. The effects of attenuation on overall performance were very similar for the two tasks, and there was no interaction between task and attenuation ($p = 0.965$). Most notably, the greater accuracy for tone discrimination than mode identification found for the larger attenuations in experiment 1 did not occur here. This indicates that replacing the pair of complex tones making up the perfect fifth (experiment 1) with a masking complex with an F0 one octave below that of the tonic was effective in preventing listeners from inferring which tone matched the embedded third using sequential comparison. In this regard, it should be acknowledged that the stimulus configuration used here did not necessarily rule out all unintended musical cues. Although it was considered an unlikely strategy, a decision involving a sequential comparison between the pure tones and the complex based on distinguishing the major and minor tenth might have been possible. There was, however, no indication in the results that listeners were able to do so.

Two interactions were significant. One was the interaction between task and stimulus mode ($p = 0.020$), which arose mainly because scores were higher overall for major chords in the mode-identification task but for minor chords in the tone-discrimination task (Fig. 3, middle vs bottom panels). The other was the three-way interaction ($p = 0.011$), which arose because the interaction of stimulus mode with task was strongest at mid-range attenuations. There is no obvious reason for these interactions involving stimulus mode, and they are not considered further.

As noted above, the psychometric functions obtained in this experiment for the mode-identification and tone-discrimination tasks were very similar. For a notional threshold corresponding to 75% performance, the difference in tone attenuation between the two tasks was small and in the opposite direction to that predicted based on precedence of the musical mode of listening (about +2 dB). Most notably, performance in both tasks approached chance for an attenuation of 35 dB. This is in marked contrast to the corresponding experiment of Hall and Pastore (1992), for which mode-identification performance remained above 70%, even at the lowest level tested (–30 dB relative to the pair of complexes making up the perfect fifth), despite chance performance in tone discrimination. Following Bailey and Herrmann (1993), this outcome indicates that comparable performance is obtained for the two kinds of judgment when stimulus design and task constraints are optimized to reduce or eliminate unintended cues and criterion effects. There was no evidence for subliminal perception.

IV. CONCLUDING REMARKS

The duplex perception of speech continues to interest researchers [e.g., Stilp *et al.* (2022) and Rizzi and Bidelman

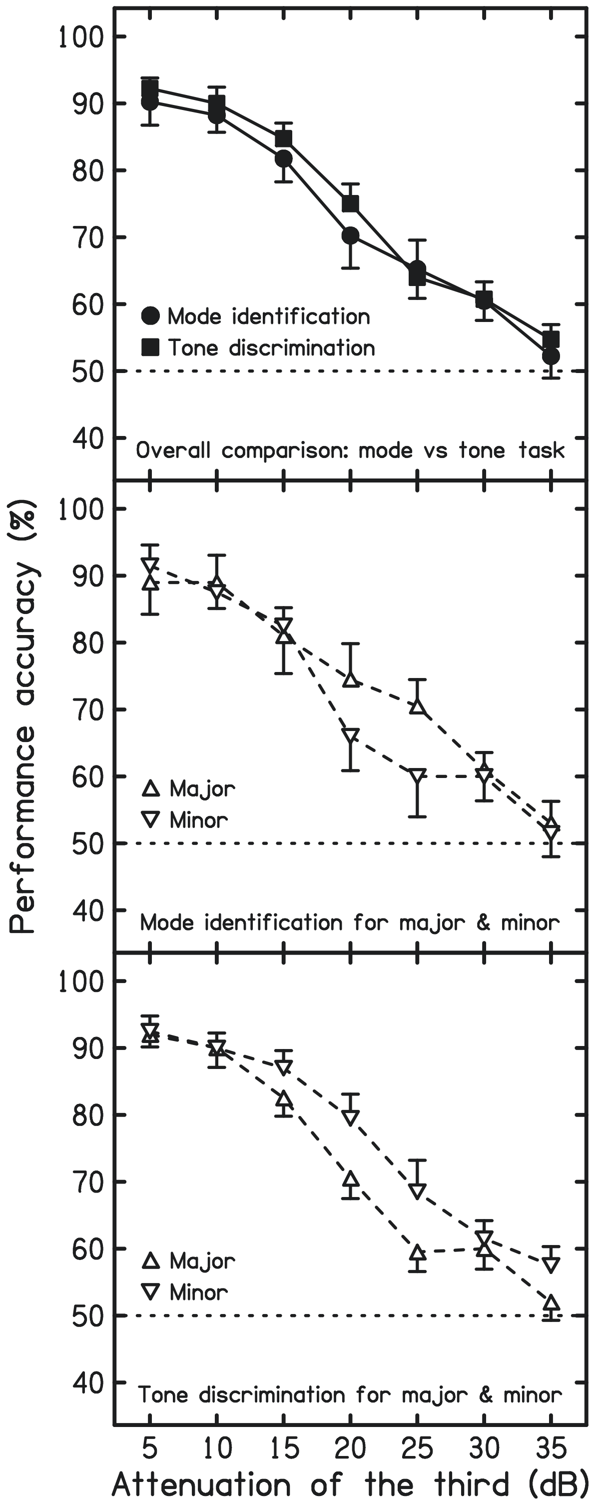

FIG. 3. Results of experiment 2. (Top panel) Results for each task, collapsed across stimulus mode. (Middle and bottom panels) Results are shown separately for major and minor stimulus configurations in the mode-identification and tone-discrimination tasks, respectively.

(2023)], but the duplex perception of musical chords has received little or no attention since the study of Hall and Pastore (1992). Introducing a difference in acoustic form between the tonic and fifth (harmonic complexes) and the third (pure tone) undoubtedly leads to duplex perception,

TABLE II. Summary of the three-way repeated-measures ANOVA for experiment 2. Task had two levels (chord identification or tone discrimination), stimulus mode had two levels (major or minor), and attenuation (intensity of the third) had seven levels (5–35 dB, in 5-dB steps). All significant terms are shown in bold.

Factor	*df*	*F*	*p*	η_p^2
Task (T)	(1, 9)	0.477	0.507	0.050
Mode (M)	(1, 9)	0.401	0.543	0.043
Attenuation (A)	**(6, 54)**	**74.330**	**<0.001**	**0.892**
T × M	**(1, 9)**	**7.947**	**0.020**	**0.469**
T × A	(6, 54)	0.231	0.965	0.025
M × A	(6, 54)	0.331	0.918	0.035
T × M × A	**(6, 54)**	**3.093**	**0.011**	**0.256**

but the experiments reported here contradict the claim by Hall and Pastore (1992) that the duplex perception of musical chords involves subliminal perception. When care is taken to reduce greatly or eliminate criterion differences between corresponding tasks, and opportunities for inference and the associative learning of unintended cues, differences in performance across mode and tone tasks are largely absent. This outcome parallels that reported by Bailey and Herrmann (1993) for the duplex perception of CV syllables and is not consistent with the notion of priority access to the sensory data (precedence) for categorical judgments of speech or music. At presentation levels where the critical sensory evidence to support discrimination is audible, synthetic and analytic modes of listening both have access to the information it carries.

The results of the study presented here, and those of Bailey and Herrmann (1993), are in accord with the account of duplex perception offered by Bregman (1987, 1990). In that account, duplex perception is considered to arise in the context of acoustical transparency—sounds originating from the same direction mix rather than occluding one another—and the low weighting given to spatial cues for auditory grouping when in competition with other cues [e.g., Deutsch (1975) and Darwin and Ciocca (1992)]. In this context, duplex perception arises when the critical acoustic element fits well into two perceptual organizations at once. For example, when the distinguishing F3 transition or third note is presented in a different ear or with different source properties to the rest of the sound, there is evidence that it forms a separate auditory object, but at the same time it completes a "strong form" based on highly learned schemas of category (speech or music), enabling a judgment of the whole sound.

ACKNOWLEDGMENTS

This research was supported by the Biotechnology and Biological Sciences Research Council (UK), which funded a Ph.D. studentship at the University of Birmingham under the supervision of B.R. I am grateful to Peter Bailey and Brian Moore for their comments on this manuscript.

https://doi.org/10.1121/10.0038976

AUTHOR DECLARATIONS

Conflict of Interest

The author has no conflicts to disclose.

Ethics Approval

Ethics approval for this research was obtained from the Psychology Research Ethics Committee at the University of Birmingham, Birmingham, UK.

DATA AVAILABILITY

The research data underlying this publication are available online from a repository hosted by Aston University at https://doi.org/10.17036/researchdata.aston.ac.uk.00000657 (last viewed on July 23, 2025).

Bailey, P. J., and Herrmann, P. (**1993**). "A reexamination of duplex perception evoked by intensity differences," Percept. Psychophys. **54**, 20–32.

Bregman, A. S. (**1987**). "The meaning of duplex perception: Sounds as transparent objects," in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Nijhoff, Dordrecht, the Netherlands), pp. 95–111.

Bregman, A. S. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT Press, Cambridge, MA).

Bruce, G. (**2000**). "An investigation of duplex perception with musical triads," Ph.D. thesis, University of Birmingham, Birmingham, UK.

Burns, E. M., and Ward, W. D. (**1982**). "Intervals, scales, and tuning," in *The Psychology of Music*, edited by D. Deutsch (Academic Press, New York), pp. 241–269.

Collins, S. C. (**1985**). "Duplex perception with musical stimuli: A further investigation," Percept. Psychophys. **38**, 172–177.

Darwin, C. J., and Ciocca, V. (**1992**). "Grouping in pitch perception: Effects of onset asynchrony and ear of presentation of a mistuned component," J. Acoust. Soc. Am. **91**, 3381–3390.

Deutsch, D. (**1975**). "Two-channel listening to musical scales," J. Acoust. Soc. Am. **57**, 1156–1160.

Fowler, C. A., and Rosenblum, L. D. (**1990**). "Duplex perception: A comparison of monosyllables and slamming doors," J. Exp. Psychol. Hum. Percept. Perform. **16**, 742–754.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103–138.

Hall, M. D., and Pastore, R. E. (**1992**). "Musical duplex perception: Perception of figurally good chords with subliminal distinguishing tones," J. Exp. Psychol. Hum. Percept. Perform. **18**, 752–762.

Henke, W. L. (**1997**). "MITSYN: A coherent family of high-level languages for time signal processing."

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Liberman, A. M. (**1982**). "On finding that speech is special," Am. Psychol. **37**, 148–167.

Liberman, A. M., and Mattingly, I. G. (**1989**). "A specialization for speech perception," Science **243**, 489–494.

Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (**1985**). "Relative dominance of individual partials in determining the pitch of complex tones," J. Acoust. Soc. Am. **77**, 1853–1860.

Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (**1986**). "Thresholds for hearing mistuned partials as separate tones in harmonic complexes," J. Acoust. Soc. Am. **80**, 479–483.

Pastore, R. E., Schmuckler, M. A., Rosenblum, L., and Szczesiul, R. (**1983**). "Duplex perception with musical stimuli," Percept. Psychophys. **33**, 469–474.

Rand, T. C. (**1974**). "Dichotic release from masking for speech," J. Acoust. Soc. Am. **55**, 678–680.

Rizzi, R., and Bidelman, G. M. (**2023**). "Duplex perception reveals brainstem auditory representations are modulated by listeners' ongoing percept for speech," Cereb. Cortex **33**, 10076–10086.

Stilp, C. E., Shorey, A. E., and King, C. J. (**2022**). "Nonspeech sounds are not all equally good at being nonspeech," J. Acoust. Soc. Am. **152**, 1842–1849.

Vos, J. (**1982**). "The perception of pure and mistuned musical fifths and major thirds: Thresholds for discrimination, beats, and identification," Percept. Psychophys. **32**, 297–313.

Whalen, D. H., and Liberman, A. M. (**1987**). "Speech perception takes precedence over nonspeech perception," Science **237**, 169–171.