

F0 と F2 の高さを同時に知覚させる歌唱法 「ホーミー」について

山田真司 杉村信彦 安東泰宏 中山一郎

大阪芸術大学音楽学科

585 大阪府南河内郡河南町東山

「ホーミー」は、1名の歌手が同時に二つの高さを表現する技法を持った、モンゴルの伝統的歌唱法である。本研究では、まず、各鍵に純音を割り当てたシンセサイザーキーボードを用いて、ホーミーの高さの特徴を模倣させる実験を行った。その結果、われわれがホーミーを聴いたときに知覚する二つの高さは、第2ホルマントと基本周波数に対応することが分かった。次に、ホーミーをシミュレートした合成音を作成し試聴した結果、第2ホルマントの共振が鋭いことと同時に、この共振によって強調された成分にゆっくりとした深い振幅変調が含まれていることを手がかりとして二つの高さが知覚されていることが分かった。

Why do we perceive two pitches which correspond to the fundamental frequency and the second formant in Xöömij sounds?

Masashi Yamada Nobuhiko Sugimura Yasuhiro Ando Ichiro Nakayama

Department of Musicology, Osaka University of Arts

Higashiyama, Kanan-Cho, Minamikawachi-Gun, Osaka 585, Japan

Xöömij is a Mongolian singing style whereby a singer generates two pitches simultaneously. Pitches in biphonic Xöömij sounds were simulated by musicians with a synthesizer keyboard which generated pure tones. Analysis of these simulated sounds confirmed that the two pitches in Xöömij sounds corresponded to the fundamental frequency and the second formant frequency. Secondly, artificial Xöömij sound were simulated by computer. These sounds indicated that biphonic perception was created not only by the sharp resonance of the second formant that increases the intensity of a particular harmonic, but also by the deep and slow fluctuation in amplitude of the harmonic.

1. はじめに

モンゴルの伝統的歌唱法「ホーミー」は、1名の歌手が同時に二つの高さの音を表現する技術を持った歌唱法として知られている。通常、1名が発する音声や歌声を聴くとき、われわれは基本周波数 (F_0) に対応する単一の高さのみを知覚する。通常の調波複合音でも、分析的な構えで聴くことにより、成分音の高さを知覚できることを Helmholtz (1859) が述べている。しかし、ホーミー音楽を聴くときには、通奏低音のようにこのような分析的な構えなしに自然に、低く一定した音高が続く中で、甲高い音をメロディーとして知覚できる。このような特殊な歌声であるホーミーに関して、過去にその音響特性について物理的に研究が行われている (Tran and Guillou, 1980; 村岡他, 1984; 武田他, 1992, 1993)。このような研究の中で、ホーミーの音において、特定の部分音が急峻な共振によって非常に強く強調されていることが示されており、このことによって、その部分音の高さが聴こえているのではないかと説明されている。この説は「共振説」と呼ばれている。しかしこれらの研究は、ホーミーの物理的音響特性が通常の歌声や音声とどのように違うかを明らかにすることに主眼がおかれており、どの高さが聴こえているのかは心理学的な実験によっては確定されていない。また、なぜそのような特殊な知覚が生じるのかについても聴覚心理学的な立場からは議論されていない。

そこで本研究では、まず第一に、われわれはホーミーを聴くとき、この音を構成する部分音の中のいずれの周波数に関係する高さを知覚するのかについて実験を行い、第二に、さまざまな合成音の試聴によって、なぜそのような知覚が生じるのかについて、Bregman (1990) の言う音情景分析 (Auditory Scene Analysis) の立場から議論する。

2. どの高さが聴こえるのか

JISによって、音の高さは「正常な聴覚の人が聞いて、それと同じ高さをもつと思われる特定の音圧の純音の周波数で表す事が出来る」とされている (JIS Z 8109-1976, 4001)。したがって、ホーミーでどの高さが聴こえるのかを評価するために、純音を発生させる装置を用いてその高さを模倣させ、この純音の周波数を測定する方法を用いることは妥当と思われる。しかし、通常の純音発生器の周波数をつまみなどで調整させる方法は、すばやく高さが推移するメロディーを持ったホーミー音楽を模倣するには適していない。そこで今回は、純音を出力するシンセサイザーキーボードを用いて、このような鍵盤楽器の扱いに習熟した音楽家に、ホーミーの高さを模倣させる実験を行うこととした。

2.1. 実験1

ホーミー音楽の CD (KING RECORD, K30Y 5118, 1988) から、伴奏がなく1名の歌手が2音を表現している箇所を2箇所採取 (9秒, 12秒) し、これをDATに録音した (以後この音を Original と呼ぶ)。一方、シンセサイザーキーボード YAMAHA DX7II の出力波形を正弦波に設定し、音楽家に Original のホーミー音楽の「高さの特徴」を、このキーボードを用いて模倣させる実験を行った。Original を呈示し、被験者にこの音楽の高さの特徴をキーボードで模倣させることを1試行とし、このような試行を被験者が納得出来る模倣を完成できるまで繰り返し、最後に納得できた模倣音楽をDATに収録した。この模倣音楽を以後、Simulated と呼ぶ。被験者は、大阪芸術大学音楽学科の教員4名と学生2名の計6名で、いずれもピアノ演奏に熟練した音楽家である。なお Original のホーミー音およびキーボー

下の出力は、ヘッドホン stax A-pro を通し、いずれも約 75 dB(A) で両耳に呈示された。また、Original を呈示する間は、キーボードの音をヘッドホンに出力しないようにして、Original の音を聴きながらさまざまな鍵をさぐることが出来ないようにした。

2.2. 実験2

実験1で求められた Simulated をそれぞれ Original と一対にして呈示し、被験者に「二つの音楽の高さの特徴」がどの程度似ているかを評定させた。被験者は音楽に習熟した大阪芸術大学音楽学科学学生 15 名であり、音源の呈示レベルは約 75 dB(A) で、スピーカを通して呈示した。評定尺度は、「全く似ていない」から「非常に似ている」の 7 段階とした。

2.3. 結果と考察

実験1で得られた、Simulated のうち、1 名のものは同時に 8 以上の鍵を押さえて表現を行っており、もう 1 名のものは、低い 1 音のみを表現していた。残りの 4 名のものは、低い音の鍵を押さえ続ける一方で、高い音でメロディーを奏でていた。この 4 名の押さえる鍵はほぼ一致していた。ただしこの 4 名の中の 1 名は低い音を他の者より 1 オクターブ低い音で表現していた。また、同時に 8 音以上の鍵を押さえたものも、これら 4 名の押さえた鍵を含み、その他は最低音の倍音系列内に含まれる鍵であったこと、および演奏した音楽家の内観報告から、この被験者はホームーの「高さの特徴」だけではなく、音色の特徴も同時に模倣したと考えられる。

実験2の評定実験の結果、低い 1 音のみを表現していた Simulated は Original とあまり似ていないと評定されたが（「全く似ていない」を 0、「非常に似ている」を 6 とし、被験者にわたる評定の平均値 0.35, 0.45）、その他の 5 名のものはいずれも同様の高い評価を得た（平均値 3.1~4.2）。Original と高さの特徴が最も似ていると評価された Simulated の狭帯域ソナグラフ、および 8 次の LPC 分析結果の一部を Fig.1 に示す。また、この部分が模倣していると思われる箇所の Original の狭帯域ソナグラフおよび 8 次の LPC 分析結果を Fig.2 に示す。Fig.1 と Fig.2 を比較すると、Simulated で表現された二つの高さのうち、上の高さは Original の第 2 ホルマント (F_2) によって強調された成分の周波数に相当することが分かる。一方、低い一定した高さの音は、第 1 ホルマントより低い周波数であることが分かる。Simulated と Original の FFT 分析の結果、この Simulated の低い音は Original の基本音の周波数に対応することが分かった。また FFT 分析の結果、ここで採用した Original のホームー音の基本周波数は 187.5 Hz で、上のメロディーのように聴こえる高さは第 2 ホルマントにより強調された第 6 倍音から第 10 倍音であることが分かった。

実験1および実験2の被験者に内観を求めたところ、ほとんどの者は、Original に含まれる低く一定した音の音色は豊かな音色であったと報告した。また多くの被験者によって、この音色は時間と共に変化するように知覚されたことが報告された。一方、メロディーを構成する高い音の音色は終始一定しており甲高く感じられたことが報告された。これらの内観報告から、第 2 ホルマント付近の成分音がその他の成分から分凝して知覚されていると考えられる。さて、実験1の被験者のうち 1 名は低い方の 1 音しか表現しなかったが、内観からも、この 1 名の被験者は高い音を知覚していないことが分かった。同じ Original の音を数名の音響研究者や音楽家に試聴させたところ、ほとんどの者は実験1の多くの被験者と同様、低い一定した高さと甲高いメロディーとを同時に知覚したが、1 名の音声研究者は、こ

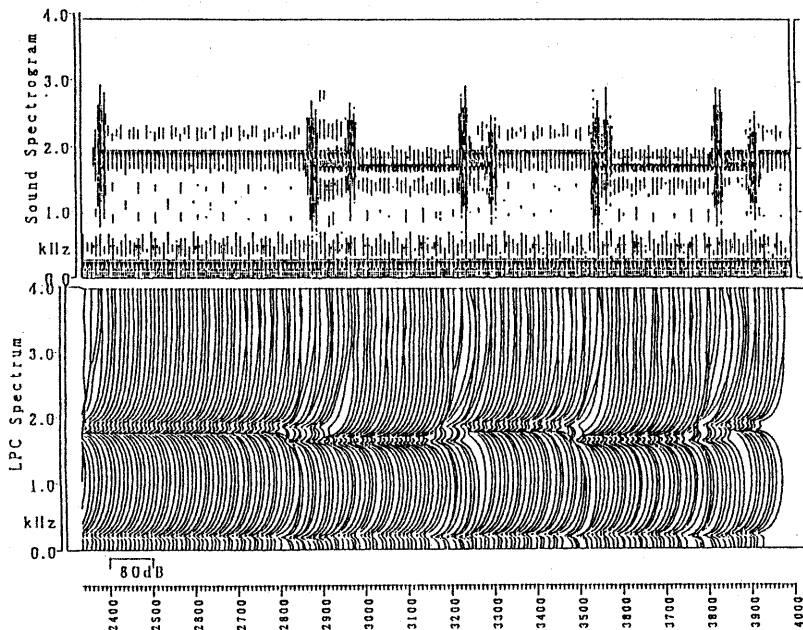


Fig. 1 各鍵に純音を割り当てたシンセサイザーキーボードを用いてホーミーの高さの特徴を模倣した音の狭帯域サウンドスペクトログラム（上段）と8次のLPC分析結果（下段）。

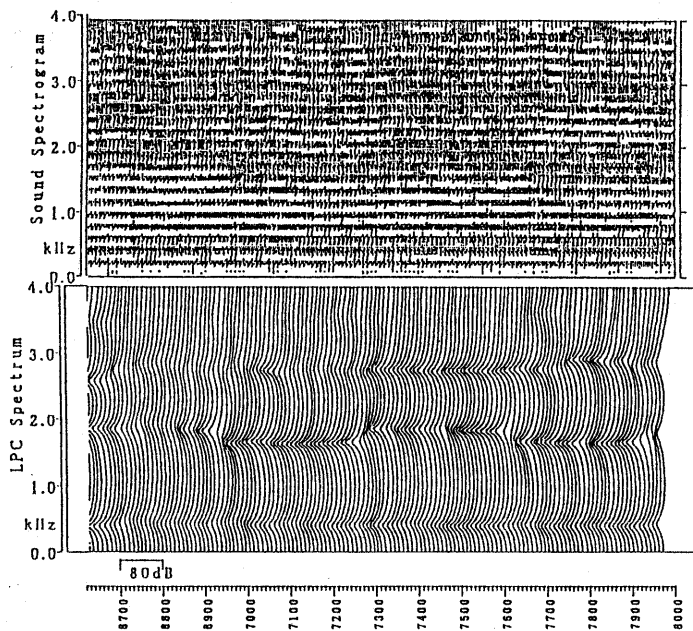


Fig. 2 オリジナルのホーミー音の狭帯域サウンドスペクトログラム（上段）と8次のLPC分析結果（下段）。Fig.1 で示した模倣音が模倣した部分を示す。

れを何度聴いても、低い一定した高さの中で音韻が変化しているようにしか知覚しなかった。これらのことから、ホーミーの音楽を聴いたときの二つの高さの知覚は、だれにでも共通に生じる原始的な分凝 (primitive segregation) の過程だけでは説明できず、スキーマを基礎とした分凝 (schema-based segregation) の過程も含まれていると考えられる。

3. なぜ2つの高さが聴こえるのか

これまでのことから、ホーミーで知覚される二つの高さは、第2ホルマントで強調された倍音成分が、他の成分から分凝することにより生じていることが示唆された。ここでは、このような分凝を行わせる物理手がかりは何なのかについて調べる。ただし、以下に述べる結果は形式的な心理実験によって確かめられたものではなく、予備実験的に数名の内観によって非形式的に示されたものである。

3.1. Old-Plus-New Heuristic で説明できるか

刺激中の何れかの成分に変化があったとき、この変化のあった成分が変化がなかった成分よりも目立って知覚されるという、われわれが経験的に知っている法則は、Old-Plus-New Heuristic と呼ばれる。Kubovy (1981) は、音階中の各音の周波数を持つ成分音で構成される非調波複合音を合成し、次々にこれらの成分の中のいずれかの強さにすばやい変化を加えることによって、メロディーを知覚させることができることを示した。Fig.2 から、第6倍音から第10倍音の中の何れかの成分が第2ホルマントの急峻な共振特性によって選択的に強調されており、この強調が起こる成分はつぎつぎとすばやく推移している様子が見てとれる。この音響的特徴は、調波複合音と非調波複合音の違いはあるが、Kubovy のデモンストレーションと共通するものである。特定の成分の強さが急に増大または減少することによってその成分が浮かび上がって知覚されることが、ホーミーにおける二つの高さの知覚の主たる原因であるとの仮説を立てられる。

この仮説を確かめるために、Fig.2 の 7280 ms の時点から 300 ms の区間を切り出し (以後「切り出し音」と呼ぶ)、これを試聴した。この切り出し音は第9倍音が第2ホルマントによって強調されている部分であり、切り出し音の中では強調された成分の推移はみられない。もし特定の成分の強さの突然の変化がホーミーにおける二つの高さの知覚の決定的な原因であるならば、このような切り出し音では、二つの高さが知覚できないはずである。しかしながら、この切り出し音を試聴した全員が二つの高さを知覚した。このことから、特定成分の突然の強さの変化が二つの高さの知覚の決定的な原因ではないことが分かる。

3.2. 成分音の強さの特徴で説明できるか

切り出し音の 300 ms にわたる長期スペクトルを Fig.3 に示す。Fig.3 から、第9倍音が周辺の成分に比して非常に強いことが分かる。これは第2ホルマントの急峻な共振特性によるものと思われる。通常の音声や歌声においては、ホルマントの共振特性はこれほど急峻ではなく、スペクトルエンベロープはなだらかである。しかし、この切り出し音の場合は第9倍音だけが他の成分からとびぬけて強く、スペクトルエンベロープはなだらかではない。このような特性は過去のホーミーの音響学的研究においても同様にみられており、この特性がホーミーの二つの高さを決定づける原因であるという「共振説」

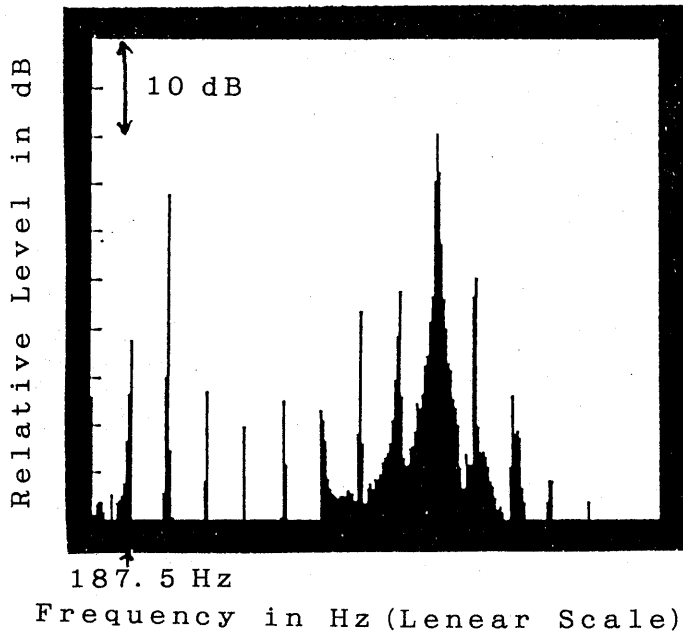


Fig. 3 ホーミー音のなかで音高が一定した箇所の長期間パワースペクトルの例。
 Fig.2 の 7280 ms ~ 7580 ms の区間を切り出し、この音のスペクトルを分析した。
 基本周波数 187.5 Hz。第2ホルモントの急峻な共振特性によって、第9倍音が強大なパワーを持っていることが分かる。

が唱えられている。この「共振説」を音情景分析の立場から考えると、自然界において、このような不連続なスペクトルエンベロープを持つ音が単一の音源から発せられることはまれであり、このことから、聴覚系がこの不連続性の原因である強大な成分を他の成分から分離させるのだと説明できる。

この仮説を確かめるための前段階として、まず、切り出し音の第11倍音以上をローパスフィルタによって除去し、この音を試聴した。その結果、試聴した全員が切り出し音同様に二つの高さを知覚した。したがって、切り出し音の基本音から第10倍音までの中に、二つの高さを知覚させる物理的手がかりが含まれていることが分かった。

次に、Fig.3の基本音から第10倍音までの各成分と同じ周波数と強さの成分から構成される合成音を作成した。このとき、Fig.3で見られる側帯波を発生させている振幅・周波数変調はつけず、ゆらぎのない定常音を作成することとした。この合成音を試聴したところ、過去に、もとの切り出し音、またはホーミーの音楽を聴いた経験のある者は、二つの高さを知覚出来るが多かったが、この合成音を最初に聴かせた場合、半数程度の者が低い単一の高さしか知覚しなかった。このことから、旧来から言われている「共振説」だけでは、ホーミーにおける二つの高さの知覚を説明できないことが分かる。

3.3. 「ゆらぎ説」の提唱

Fig.3から第9倍音とその周辺の成分には側帯波が含まれていることが分かる。また Fig.2より第2ホルモント付近にはゆっくりとした深い振幅変調が含まれていることが観測される。そこで、切り出し音の基本音から第10倍音までを 135 dB/Oct. の急峻な特性を持つバンドパスフィルターによって分

離し、各成分毎に周波数変調および振幅変調の様子を分析した。その結果、何れの成分音においても周波数変調はほとんど含まれていないことが分かった。一方、振幅変調の様子は、成分によってかなり異なっていた。基本音から第 6 倍音までにはほとんど振幅変調は含まれていなかった。しかしながら、第 9 倍音には、変調周波数約 2 Hz の変調が約 25 % の深さで含まれていた。またその周辺の倍音にはこれほど深さは深くはないものの、第 9 倍音の変調に同期した振幅変調が含まれていた。

複合音中のいくつかの成分に同期した振幅変調をかけると、これらの成分は他の成分から分凝しやすいことが、Bregman et al. (1985), Moore (1989) によって示され、説明されている。このことは、旧来からゲシュタルト心理学者達が述べている「共通運命の原理」の説明と同様のものである。つまり、同じ音源から発せられる成分音は同期した動きをするはずであり、逆に異なる音源から発せられる成分の間で同期した動きが長時間観測されることはまれである。この事実を用いて、われわれの聴覚系は各成分の動きの同期性を分析し、うまく各成分音を本来のグループに分凝していると説明できるのである。

ひるがえって、第 9 倍音を中心とした第 2 ホルマント付近には、他の成分にはみられない大きな振幅変調が含まれていることを考えると、このような振幅変調の特徴が原因で、第 9 倍音付近の成分を他の成分と「共通運命」を有さず分凝して知覚されるのではないかと考えられる。この仮説を調べるために、先ほど作成した合成音の第 8 ～第 10 倍音に、もとの切り出し音から抽出したゆっくりとした深い振幅変調をかけて合成を行った。この合成音を試聴したところ、ゆらぎの無い合成音を聴いたとき低い単一の高さしか知覚しなかった者も含め、ほとんどの者が二つの高さを知覚した。このことから、第 2 ホルマント付近の振幅変調を、ホーミーにおける二つの高さの知覚の原因となる物理的手がかりのひとつとして挙げる事が出来るであろう。ここに、われわれは、ホーミーにおける特殊な知覚の原因として、「共振説」の他に「ゆらぎ説（振幅変調説）」を提唱する。

3.4. 第 1 ホルマントでもメロディーを表現できるのか

欠けた基本音 (missing fundamental) の複数の低次倍音からなる複合音の、構成音全てに同期した周波数変調を加えると、これらの成分は融合しやすくなり、欠けた基本音の高さが知覚されやすくなることが示されている (Thurlow and Small, 1955; Plomp, 1976)。このことをふまえて、現実には生理学的に可能かどうかは度外視して、ホーミーのメロディーが第 2 ホルマントではなく第 1 ホルマントによって表現された場合のことを考えてみよう。またこのとき第 1 ホルマント付近に、現実のホーミーの第 2 ホルマント付近と同様な振幅変調が含まれていたとしよう。この場合、低次倍音のそれぞれは異なる臨界帯域に含まれ、それら異なるチャンネル間で同期した動きを持つ共通運命の成分は統合されるであろう。このことによって、これらの成分に共通する基本音の高さしか知覚されないであろう。一方、現実のホーミー音楽は第 2 ホルマントを使って比較的高次の倍音でメロディーが表現されているため、近隣の倍音成分に同期した変調が含まれていても、これらの近隣の倍音は同一臨界帯域内に含まれてしまい、これらの成分を聴覚系は分離できない。したがって、この臨界帯域のチャンネルのみが他のチャンネルと異なる動きをもつためこのチャンネルだけが分凝し、独立した高さを産出すると考えられる。この場合、この分凝されたチャンネルが産出する高さは、チャンネル内で最も強い共振成分の周波数によって支配されるのであろうと考えられる。このように考えると、現実のホーミーにおいて第 1 ホルマントではなく第 2 ホルマントによって、倍音成分を選択的に鋭く共振させメロディーを表現していることは、発声の機構にとっての意味を除いても、知覚だけにとっても深い意味を持つと思われる。

4. まとめ

以上のように、ホーミーは、聴覚系の音情景分析の過程を逆手に取って、うまく1名の歌手が発した音を二つの流れに分凝させて知覚させることに成功していると考えられる。今回、このような分凝の手がかりとなる物理的特性として「ゆらぎ説」を提唱したが、決して、「共振説」や Old-Plus-New Heuristic の関与を否定しているわけではない。これらもあいまってホーミーの特殊な知覚が形成されていると考えるのが妥当であろう。今後形式的な心理実験によって、これらの要因がそれぞれ、どの程度ホーミーの特殊な知覚に寄与しているのかについて検討して行く。

[謝辞]

本研究に関してご議論いただいた、河原英紀、津崎実、加藤宏明、倉片憲治、松井理直の各氏をはじめとする、A. S. Bregman's 'Auditory Scene Analysis' 輪読会のメンバー各位に感謝します。また、データ処理に協力いただいた、渡邊守氏、杉村清美氏をはじめ、大阪芸術大学音楽工学サブスタジオの皆様にも感謝します。

[文献]

- Bregman, A.S.(1990). Auditory Scene Analysis, London: The MIT Press.
- Bregman, A.S., Abramson, J., Doehring, p., and Darwin, C.J.(1985). Spectral Integration based on common amplitude modelation. Perception & Psychophysics, 13, 451-454.
- Helmholtz, H. von (1859). On the Sensation of Tone as a Physiological Basis for the Theory of Music. (2nd English ed. Translated by A.J. Ellis, 1885). Reprinted by Dover Publications, 1954.
- Kubovy, M.(1981). Concurrent-pitch segregation and the theory of indispensable attributes. In M.Kubovy and J.R.Pomerantz(eds.), Perceptual Organization. Hillsdale, N.J.: Erlbaum. pp.66-69.
- 武田昌一, 糸賀昌士, 佐藤雄樹, 上田康夫(1992). モンゴルの歌唱法「ホーミー」の音響的特徴の解析, 日本音響学会秋季講演論文集, pp.605-606.
- 武田昌一, 糸賀昌士(1993). モンゴルの歌唱法「ホーミー」の音韻および音高の違いによるスペクトルの差異について, 日本音響学会春季講演論文集, pp.499-450.
- Tran, Q.H. and Guillou, D.(1980). Original research and acoustical analysis in connection with the Xoomij style of biphonic singing, Musical Voices of Asia, Tokyo: Heibonsha. pp.162-173.
- Moore, B.C.J.(1989). An Introduction to Psychology of Hearing (3rd ed.), London: Academic Press. pp.239-240.
- 村岡輝雄, 我妻幾久寿, 土金由幸, 堀内正人(1984). モンゴル歌唱「ホーミー」とその特徴について, 音楽音響研究会資料 MA84-1.
- CD: ホーミーの音(1988), モンゴルしょうようホーミーとオルテンドー所有 King Record co., K30Y 5118.