

聴覚障害者の環境音同定における音響特徴の考察

湯野 悠希¹ 松原 正樹² 田原 敬³ 寺澤 洋子² 平賀 瑠美⁴

概要：聴覚障害者が音を聴取する際、聴覚障害の影響で聞き取れる周波数の範囲が狭くなり、周波数の解像度が低くなることに加えて音の歪みも発生する。そのため、聴覚障害者が聴取している音は、健聴者が聴取している音とは大きく異なっていると推測される。そこで本研究では、聴覚障害者が環境音の名称を判断する際に影響を及ぼす音響特徴量の検討を目的とした。音響特徴量については先行研究を参考にして、周波数情報・振幅情報についての統計量や、自己相関を用いた時間変化などの分析を行なった。同定課題は、大学生・大学院生の聴覚障害者 15 名を対象とし、環境音を聴取して名称を自由記述で解答してもらった。以上の結果を基に、同定課題の正答率の結果と、課題で使用した刺激音の音響特徴量の関連を考察する。

YUNO YUUKI¹ MATSUBARA MASAKI² TABARU KEI³ TERASAWA HIROKO² HIRAGA RUMI⁴

1. はじめに

我々は、日常的に様々な音を聴取しながら生活している。Gygi らによれば、音は音楽・音声・環境音の 3 つに分けることが可能であり [1]、本研究では環境音を「音楽・音声を除いた日常的に聴取する音」と定義する。環境音は、日常生活において重要な役割を担っている。その役割とは大きく危険の察知と QoL の向上の 2 つに分けることが可能である [2], [3]。

近年、補聴器や人工内耳といった聴覚補償機器の性能が向上したことによって、聴覚障害者はより多くの音を聴取できるようになった。そのため、環境音の知覚・認知についても目が向けられるようになった [4], [3]。一方で、人工内耳に期待することとして環境音の聴取を挙げる聴覚障害者も存在し [2]、聴覚障害者自身における環境音への関心も高まっていると考えられる。以上のような現状と環境音の持つ役割を考慮すると、聴覚障害者における環境音の知覚や認知について研究することは重要であると言える。

環境音の知覚・認知には複数の段階が考えられるが、本研究では特に環境音の同定に着目した。同定とは、聞いた音に対してラベル付けを行うことである [5]。同定には、音

自体における音響特徴量 (周波数情報、振幅情報)、背景情報 (音の発生する仕組み、聴取状況、音源のカテゴリなど)、経験 (生育歴、家庭環境) などが影響すると考えられる。このうち、我々は音響特徴量に着目した。聴覚障害者の同定に影響する特徴量が具体的に数値化可能であれば、医療分野・教育分野などに幅広く応用が可能であるからである。

同定に影響する環境音の音響的な特徴として、いくつかの特徴量が示唆されている [6], [7], [8], [9]。周波数情報については、周波数のピークが低い周波数帯域にある音、あるいは低い周波数帯域のパワーが強い音が同定しやすく、振幅情報については、周期的もしくは特異的なリズムパターンを有する断続音が同定しやすいことが示されている。しかしながら、音声や音楽の研究に比べると研究の数は少なく、また対象とされる聴覚障害者も人工内耳装用者が主であるため、更なる検討が望まれる。

そこで本研究では、聴覚障害者 15 名を対象に実施した環境音の同定課題の結果と、課題で使用した刺激音の音響特徴量の関係を検討することを目的とする。

2. 環境音の聴取実験

2.1 実験参加者

実験には、15 名の聴覚障害者が参加した。15 名全員が先天性の感音難聴であり、補聴器を使用していた。4 分法を用いて平均聴力レベルを算出した結果、全実験参加者の平均聴力レベルは 67~109 dBHL であった。

¹ 筑波大学大学院図書館情報メディア研究科
図書館情報メディア専攻博士前期課程

² 筑波大学図書館情報メディア系
茨城県つくば市春日 1-2

³ 茨城大学教育学部
茨城県水戸市文京 2-1-1

⁴ 筑波技術大学産業技術学部
茨城県つくば市久保 4-3-15

表 1 使用した刺激音の名称とカテゴリ

名称	カテゴリ
野菜を切る音, タイピングの音, 足音	自然の音
雷, 雨風	家の中の音
救急車のサイレン, 車の走行音	乗り物の音
小鳥のさえずり, 犬の鳴き声	動物の音

2.2 刺激音

刺激音は、市販のデータベース *1 (サンプリング周波数: 44.1 kHz, 16 ビットリニア量子化, モノラル) に収録されている音源から選定した。

Finitzo-Hieber ら [10] の研究を参考に、「自然の音」、「家の中の音」、「乗り物の音」、「動物の音」の4つのカテゴリを満たし、日常生活で聴取する機会が多く具体的な名称を答えやすいものであることを条件として、最終的に9つの刺激音を選定した。表 1 に実験で使用した9つの刺激音の名称とカテゴリを示す。

2.3 実験環境

実験は筑波大学の防音室で行った。刺激音はパーソナルコンピュータ (Apple 社製, Mac Book Pro) に床から高さ 0.95 m の位置に置かれたスピーカ 1 台 (GENELEC 社製, 8020CPM) をつないで呈示した。実験参加者はスピーカーの正面に座り、スピーカーの中心と実験参加者の耳までの距離は 1.2 m であった。

刺激音の呈示音圧は、実験開始前に実験参加者に自由に調節してもらった。実験中は、音圧が大きすぎて不快であるといった場合を除き、変更をしないよう伝えた。実験中、音圧の調節を行った実験参加者はいなかった。調整後の呈示音圧を実験参加者の左耳の近くで測定した結果、52~72 dB SPL (A 特性) であった。

2.4 実験手続き

実験参加者は 2 秒の刺激音を 1 度だけ聴取し、音の名称もしくは擬音語を紙に自由記述した。刺激音が知覚できなかった場合は、「聞こえない」と解答した。また、同定課題の後には、解答の内容や刺激音についてのインタビューも実施した。

2.5 実験結果

聴覚障害者による同定課題の結果から、刺激音ごとに正解した人数の割合を算出したものを図 1 に示す。犬の鳴き声が最も正答率が高く、100% であった。一方、最も正答率が低かったのは、タイピング音、雷、足音の3つであり、12.5% であった。また、刺激音全体の平均正答率は 35.5% であった。

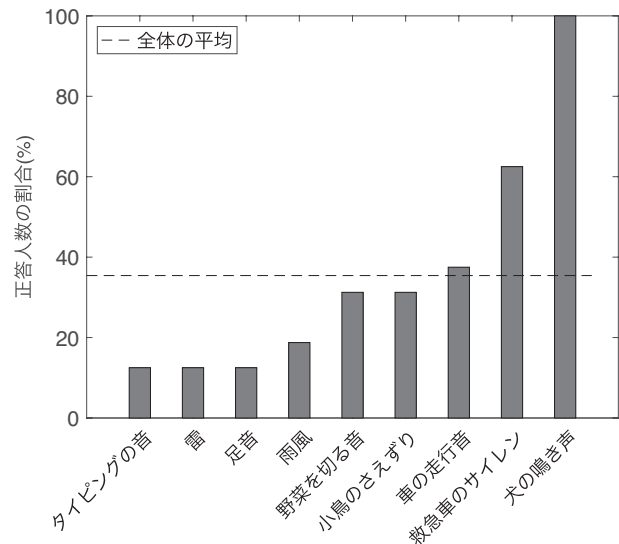


図 1 刺激音ごとにみた正答人数の割合

3. 音響特徴量

音響分析には、先行研究 [2], [11], [12], [13] を参考に 23 個の音響特徴量を使用した。表 2 にその一覧を示す。

3.1 エンベロープに関する特徴

エンベロープに関する特徴量は、波形のおおまかな時間変化を表す。エンベロープは、幅が 20 ms の矩形窓を 10 ms ずつずらしながら計算した RMS に基づいて算出した。

無音を除去した RMS とは、Ludvigsen ら [14] が用いた特徴量であり、50 ms 以上継続する無音区間を除去して算出したものである。全区間の RMS における無音を除去した RMS の割合は刺激音における無音区間の割合を示し、値が小さいほど無音区間が長いことを示す。

また、バーストとは、Ballas[11] が用いた特徴量であり、20 ms の間に振幅が 4 dB 以上増加した区間と定義されている。全区間におけるバーストの割合はエンベロープの粗さに関する特徴量であり、値が小さいほど滑らかで時間変化が少ないことを示す。

3.2 波形の自己相関に関する特徴

波形の周期性を表す特徴量として、自己相関に基づく特徴量を用いた。自己相関とは、元の信号とそれ自身を時間シフトさせた信号との相関を取ったものである。

3.3 ピッチに関する特徴

ピッチとピッチサリエンスの算出には、Slaney が開発した MATLAB の toolbox である Auditory toolbox[15] を利用した。この toolbox では自己相関を利用してピッチやピッチサリエンスを算出しており、窓幅などが指定できる。今回、ピッチの算出には 16 ms の幅の矩形窓を用いた。

*1 「新効果音大全集」K30X5004~6, 5009,
「著作権フリーデジタル音素材集音・辞典」HR-AJ03, 06,08

表 2 音響分析に使用した音響特徴量の一覧

エンベロープに関する特徴
全区間の RMS に対する無音を除去した RMS の割合
エンベロープにおけるピークの数
エンベロープにおけるバーストの数
エンベロープの全区間に対するバースト区間の割合
波形の自己相関に関する特徴
自己相関におけるピークの数
自己相関の最大値, 平均値, 標準偏差
ピッチに関する特徴
ピッチの平均値, 中央値, 標準偏差, 最大値
ピッチサリエンスの平均値, 最大値
全区間の周波数スペクトルに関する特徴
周波数スペクトルの重心, 標準偏差, 歪度, 尖度
周波数の時間変化に関する特徴
各区間ごとの周波数重心の平均値, 標準偏差
周波数重心の速度の平均値, 標準偏差, 最大値

また, ピッチサリエンスは, 1 に近いほどピッチを知覚しやすい周期性の高い音であり, 0 に近いほど周期性が低い音であることを表す [15].

3.4 全区間の周波数スペクトルに関する特徴

これらの値は, 周波数スペクトルにおける分布の特徴を表す値である. 歪度は分布の非対称性を表す指標であり, 歪度が正のとき分布は左に偏り, 0 のときは左右対称, 負のときは右に偏った分布となる. 尖度はある分布と正規分布を比べたときの分布の尖り具合を表す指標であり, 尖度が大きいと分布はより鋭い形になり, 尖度が小さいと分布はより鈍い形になる.

3.5 周波数の時間変化に関する特徴

周波数スペクトルの時間変化は, 幅が 20 ms のハミング窓を 10 ms ずつずらしながら 882 点のフーリエ変換を行うことで得られた. 周波数重心は 20 ms の区間ごとに計算され, 速度はある 20 ms の区間と次の 20 ms 区間の周波数重心の変化量によって算出した.

4. 音響分析の結果

2 章で述べた聴取実験によって得られた同定課題の正答率と, 3 章で述べた音響特徴量を用いて重回帰分析を行なった. 音響特徴量の平均が 0, 分散が 1 になるように標準化を行なった後, ステップワイズ法による特徴量の選択と重回帰分析を実施した. ステップワイズ法では, 特徴量を追加した場合としない場合で F 検定を実施し, p 値を基準として特徴量の追加と削除を行いながら最終的に特徴量を絞り込む.

その結果, 周波数スペクトルの尖度, 周波数スペクトルの歪度, 周波数重心の速度の標準偏差, 周波数重心の標準偏差, 全区間に対するバースト区間の割合, 周波数重心の

表 3 重回帰分析の結果得られた特徴量の内訳と標準化偏回帰係数 (自由度調整済み決定係数 $R^2 = 1.0$)

特徴量	標準化偏回帰係数
周波数スペクトルの尖度	33.69
周波数スペクトルの歪度	-33.43
周波数重心の速度の標準偏差	-16.67
周波数重心の標準偏差	15.94
バースト区間の割合	4.15
周波数重心の平均	-3.30
バースト区間の割合 × 周波数重心の平均	-1.19

表 4 刺激音に対する音響特徴量の値

名称	正答率	スペクトルの尖度	スペクトルの歪度	バースト区間の割合
タイピングの音	12.50	8.41	97.66	18.26
車の走行音	37.50	134.54	22912.04	0.00
犬の鳴き声	100	19.18	472.37	7.01

平均の 6 つの特徴量が残り, 全区間に対するバースト区間の割合と周波数重心の平均との間に交互作用が見られた. 表 3 に重回帰分析の結果として, 標準化偏回帰係数の絶対値が大きい順に示す. また, 9 つの刺激音における各特徴量の値の内訳を付録に記載した.

5. 考察

音響分析の結果得られた音響特徴量の中から, 今回は周波数情報として周波数スペクトルの歪度・尖度と, 振幅情報としてバースト区間の割合に着目した. また, 9 つの刺激音の中から持続音である車の走行音と, 時間変化が急な断続音であるタイピングの音と, 時間変化の緩やかな断続音である犬の鳴き声の 3 音について考察を行う. 表 4 に, 3 つの刺激音における周波数スペクトルの歪度, 尖度, バースト区間の割合についての具体的な数値を示した.

5.1 全区間に対するバースト区間の割合

車の走行音, 犬の鳴き声, タイピングの音について, 各刺激音のエンベロープとバースト区間と判定された区間を図 2 に示した. 灰色の部分が一時的にバーストと判定された区間である. 図 2 を見ると, 車の走行音は短時間の振幅変動が見られず, バースト区間の割合は 0% であるのに対し, 犬の鳴き声やタイピングの音などは短時間での振幅変動を伴うため, バースト区間の割合が車の走行音に比べて高くなっている.

先行研究によれば, 持続音よりも断続音の方が同定しやすいことが示唆されている. すなわち, バースト区間の割合が高い音は断続音に近く同定しやすいと考えられる. タイピングの音が最も正答率が高く, 車の走行音が最も正答率が低くなることが予想される. しかし, 図 1 に示した実際の同定課題の結果を見ると, 犬の鳴き声と車の走行音については断続音に近い犬の鳴き声の方が正答率が高くなっているが, 車の走行音とタイピングの音では持続音である車の走行音の方が正答率が高いことが分かる. この点について

ては、5.3 節で考察する。

5.2 周波数スペクトルの歪度・尖度

車の走行音、犬の鳴き声、タイピングの音の周波数スペクトルを図 3 に示した。この中で、車の走行音が最も歪度が高く、タイピングの音が最も歪度が低い。すなわち、車の走行音の周波数スペクトルは他の 2 音よりも分布が低い周波数に偏っていることが分かる。また、尖度に関しても、最も大きいのは車の走行音であり、最も小さいのはタイピングの音である。これらのことから、車の走行音は低い周波数帯域に分布が偏っており、さらに尖度も高いことから低い周波数帯域にパワーが集中していると考えられる。また、タイピングの音は分布の偏りが車の走行音よりも小さく、なだらかな分布であることが言える。

実際に図 3 を見ても、車の走行音の周波数は低い周波数帯域に偏っていることが分かる。一方、タイピングの音は非常に短時間で振幅が変動するクリック音に近い音であり、他 2 つの音よりも多くの周波数成分を含んでいることが読み取れる。

先行研究によれば、低い周波数帯域のパワーが強い音は同定しやすいと示唆されている。そのため、車の走行音が最も正答率が高く、タイピングの音が最も正答率が低くなることが期待される。しかし、図 1 に示した結果では犬の鳴き声が最も正答率が高く、次いで車の走行音、タイピングの音の順となっており、犬の鳴き声は期待に反する結果となった。この点について、5.3 節で考察を行う。

5.3 音響特徴量以外の要因

今回、同定課題の正答率と音響特徴量（バースト区間の割合、周波数スペクトルの歪度・尖度）との関係を考察した。その結果、得られた特徴量と正答率の関係は単純に先行研究の結果を支持するわけではなかった。その理由として、1 章で述べたように背景情報や経験といった、音響特徴量以外の要因による影響が考えられる。特に、今回は刺激音が 9 音と少なかったため、各刺激音に対する親密さなどの影響が強く出やすかった可能性がある。

6. まとめと今後の展望

今回は、聴覚障害者を対象とした環境音同定課題の正答率と刺激音の音響特徴量の関連を考察することを目的とした。複数の先行研究で用いられていた 23 個の音響特徴量を用い、ステップワイズ法による変数の選択と重回帰分析を行った結果、6 つの特徴量が残った。今回は車の走行音、犬の鳴き声、タイピングの音の 3 音に対して、残った 6 つの特徴量の中から、周波数スペクトルの歪度・尖度、全区間に対するバースト区間の割合について考察を行なった。

その結果、部分的には先行研究で示唆されている結果に当てはまるケースも見られたが、当てはまらないケースも

見られた。その理由としては、背景情報や聴取経験といった要因が、音響特徴量よりも強く影響していた可能性が考えられる。

今回の報告はあくまで予備調査の段階であるため、今後は刺激音の数を増やして同様の課題と音響分析を実施したいと考えている。また、刺激音の選定の際には、親密さや背景情報などの音響特徴量以外の要因についての統制を行う必要があると考える。

謝辞 本研究に協力してくださった聴覚障害学生の皆様へ深く感謝いたします。本研究は JSPS 科研費 26282001、17K14058 の助成を受けたものです。

参考文献

- [1] Gygi, B. and Shafiro, V.: Environmental sound research as it stands today, *Meetings on Acoustics. Acoustical Society of America* (2007).
- [2] Shafiro, V., Gygi, B., Cheng, M. Y., Vachhani, J. and Mulvey, M.: Perception of Environmental Sounds by Experienced Cochlear Implant Patients, *Ear and Hearing*, Vol. 32, No. 4, pp. 511–523 (2011).
- [3] 田原 敬, 原島恒夫, 小林優子, 堅田明義: 難聴者の環境音認知に関する研究の展望, *障害科学研究*, Vol. 26, pp. 187–196 (2012).
- [4] Shafiro, V.: Development of a large-item environmental sound test and the effects of short-term training with spectrally degraded stimuli, *Ear and Hearing*, Vol. 29, No. 5, pp. 775–790 (2008).
- [5] Erber, N. P. and Hirsh, I. J.: Auditory training, *Hearing and Deafness Fourth Edition* (Davis, H. e. a., ed.), Holt Rinehart and Winston, pp. 358–374 (1978).
- [6] 中川辰雄: 聴覚障害学生の環境音認知, *横浜国立大学教育人間科学部紀要*, I, 教育科学, Vol. 1, pp. 81–88 (1998).
- [7] 田原敬ら, 小林優子, 原島恒夫: 聴覚障害者の環境音認知に影響を及ぼす要因に関する研究, *聴覚言語障害*, Vol. 44, No. 1, pp. 43–50 (2015).
- [8] 高橋信雄, 今井秀雄: 聴覚障害児用の環境音受聴テストの試行 (その 2), *Audiology Japan*, Vol. 25, pp. 415–416 (1982).
- [9] Reed, C. M. and Delhorne, L. A.: Reception of environmental sounds through cochlear implants, *Ear and Hearing*, Vol. 26, No. 1, pp. 48–61 (2005).
- [10] Finitzo-Hieber, T., Gerling, I. J., Matkin, N. D. and Cherow-Skalka, E.: A sound effects recognition test for the pediatric audiological evaluation, *Ear and Hearing*, Vol. 1, No. 5, pp. 271–276 (1980).
- [11] Ballas, J. A.: Common factors in the identification of an assortment of brief everyday sounds, *Journal of experimental psychology: human perception and performance*, Vol. 19, No. 2, pp. 250–267 (1993).
- [12] Shafiro, V.: Identification of environmental sounds with varying spectral resolution, *Ear and Hearing*, Vol. 29, pp. 401–420 (2008).
- [13] Gygi, B., Kidd, G. R. and Watson, C. S.: Spectral-temporal factors in the identification of environmental sounds, *Journal of Acoustical Society of America*, Vol. 115, No. 3, pp. 1252–65 (2004).
- [14] Ludvigsen, C.: Comparison of certain measures of speech and noise level, *Scandinavia Audiology*, Vol. 21, pp. 23–29 (1992).
- [15] Slaney, M.: Auditory Toolbox: A MATLAB toolbox for

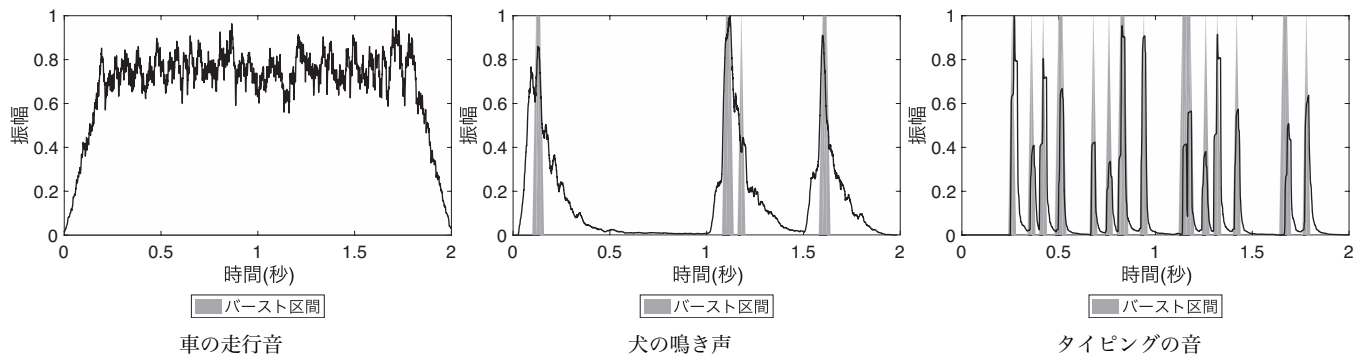


図 2 各刺激音におけるバースト区間

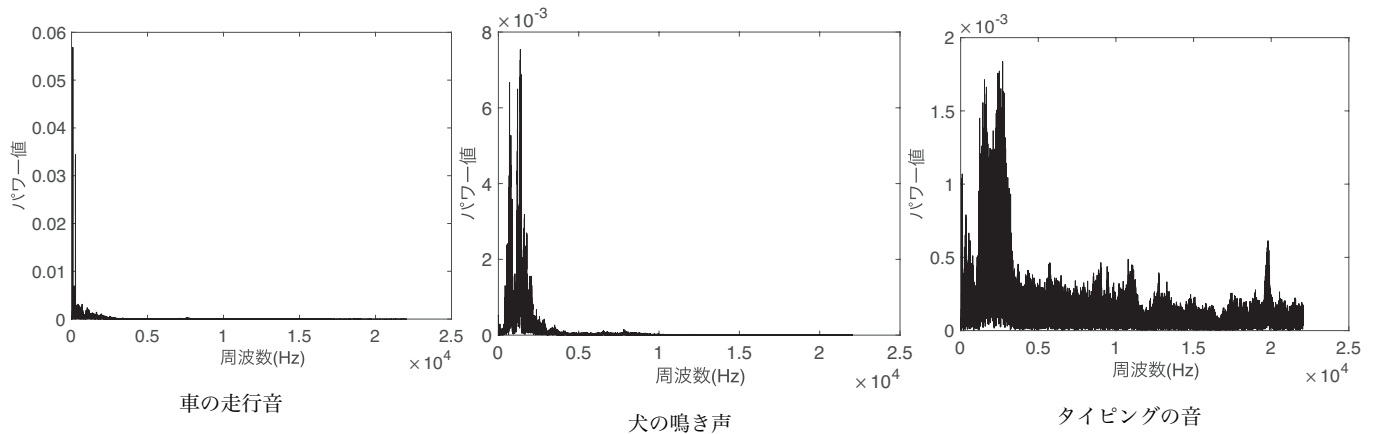


図 3 各刺激音における周波数スペクトル

auditory modeling work, Technical report, Apple Computer Technical Report #45 (1995).

付 録

A.1 各刺激音における音響特徴量の内訳

A.2 9つの刺激音の波形

表 A.1 各刺激音における音響特徴量の内訳

名称	スペクトルの尖度	スペクトルの歪度	周波数重心の標準偏差	周波数重心の速度の標準偏差	バースト区間の割合	周波数重心の平均
野菜を切る音	29.84	1146.81	11508.10	745.43	22.00	1411.36
タイピングの音	8.41	97.66	49161.22	5190.28	18.26	4865.15
足音	29.04	1234.26	4392.35	299.15	10.00	637.15
雷	21.26	668.41	622.71	48.64	2.00	343.74
雨風	11.46	262.98	2471.33	214.14	0.00	2184.52
救急車のサイレン	163.00	28780.73	294.08	77.69	0.00	933.89
車の走行音	134.54	22912.04	245.98	13.16	0.00	190.98
小鳥のさえずり	9.03	114.96	7606.61	750.80	6.00	3741.96
犬の鳴き声	19.18	472.37	753.73	79.00	7.01	788.56

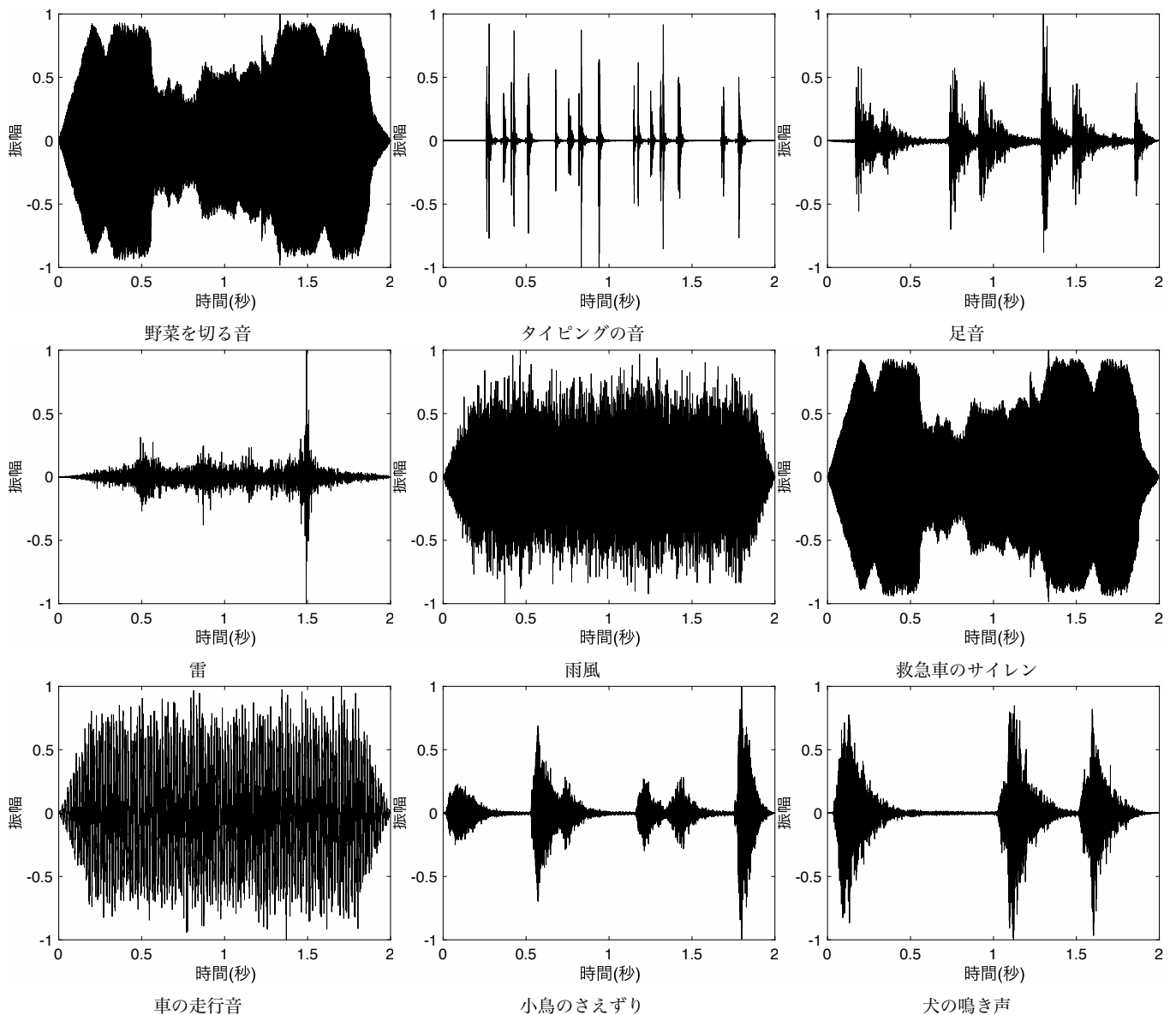


図 A.1 各刺激音の波形