



超音波を用いた調音の指導と研究*

イアン ウィルソン (会津大学語学研究センター)**・
 全訳：荒井 隆行, 石田 真子 (上智大学)***・
 溝口 愛 (ニューヨーク市立大学)****

43.70.Aj, Jt, Kv

1. はじめに

発話中の舌の動きを安全に観察し、正確に記述することは長年の課題となっている [1]。発話中の舌を観察する方法は何十年にもわたって存在していたものの、それらの多くは、高価である（核磁気共鳴画像法, magnetic resonance imaging, 略して MRI）、人体への影響が懸念される（X線による写真・動画・コンピュータ断層撮影）、扱いにくい（electromagnetic articulometry, 略して EMA）、あるいは、取得できるデータに限られる（electropalatography, 略して EPG）など、様々な理由から、理想的とはいえないものであった。しかしながら、超音波では、より安価に、安全に、そして簡単に、発話中の舌全体の画像をリアルタイムで観察することができる。超音波診断装置が発話中の舌を観察するツールとして発展し続けてきたことによって、音声研究や発音指導にも貢献している。超音波を使用した音声研究の概説は [2] を参照されたい。

超音波は 1960 年代後半には音声研究に使用されていた [3]。しかしながら、初期の研究における超音波は、レーザーポインタのような 1 次元のビームであり、2 次元の場合でも、プローブの有効視野角が狭く、舌全体をはっきりと見ることができなかった。それから約 50 年が経過し、今や、3 次元の超音波機器が存在している。しかし、発音指導や音声研究に最も一般的に用いられるのは、2 次元の超音波機器である。というのも、2 次元の画

面に映し出された 3 次元画像の解釈・測定は、しばしば困難を伴うからである。今日の音声研究に使用されるほとんどのプローブの有効視野角は、90° から 150° 超の範囲に及び、舌全体、つまり舌尖から舌根を観察することが可能となっている。

超音波診断装置で用いられる超音波の周波数（3~16 MHz）は、皮膚、脂肪、筋肉を容易に通過するが、骨には吸収され、空気層との境界面では反射する。舌を最適に観察するためには、音波が下から舌を貫通するよう、プローブは顎の下に、首に沿って保持されなければならない（図-1 参照）。下顎に対して十分に幅の狭いプローブを用い、頸部に上方へのわずかな圧力を加えることで、音波は下顎と舌骨の間を阻害されることなく通過し、舌の筋肉を通り、舌表面の空気層で反射する。プローブの向きによって、正中矢状面（midsagittal）あるいは冠状面（coronal）の画像を得ることができる。舌の冠状面超音波画像を用いた臨床研究については、[4] を参照されたい。超音波を用いて得られる舌の正中矢状面の代表的な画像は図-2 のとおりである。図-2 では、舌尖が右側であることに注意されたい。

前述のように、超音波は設定や操作が容易である。プローブを手を持ち首に当てるだけで、発話中の舌の形と動きの明瞭な画像を得ることができる。それは、舌の動きの定性的評価に十分利用可能である。ほとんどの発音指導の現場では、これだけで十分である。しかし、研究目的で定量的計測データを収集する場合など、より正確性が必要であれば、頭部に対するプローブの相対位置が動かないように注意しなければならない。プローブを金属のスタンドに保持し、後頭部を壁にもたれさせ、制御された状態でデータ収集をする場合でも、頭部は前後方向に平均 1~4 mm、上下方向に 4~9 mm 動く [5]。頭部とプローブの位置が互いに相対的に動くと、画面上の画像も動くため、舌が

* Using ultrasound for teaching and researching articulation.

** Ian Wilson (Center for Language Research, Department of Computer Science and Engineering, University of Aizu, Aizu-Wakamatsu, 965-8580)

*** Translation: Takayuki Arai and Mako Ishida (Sophia University, Tokyo, 102-8554) e-mail: arai@sophia.ac.jp

**** Ai Mizoguchi (The City University of New York)

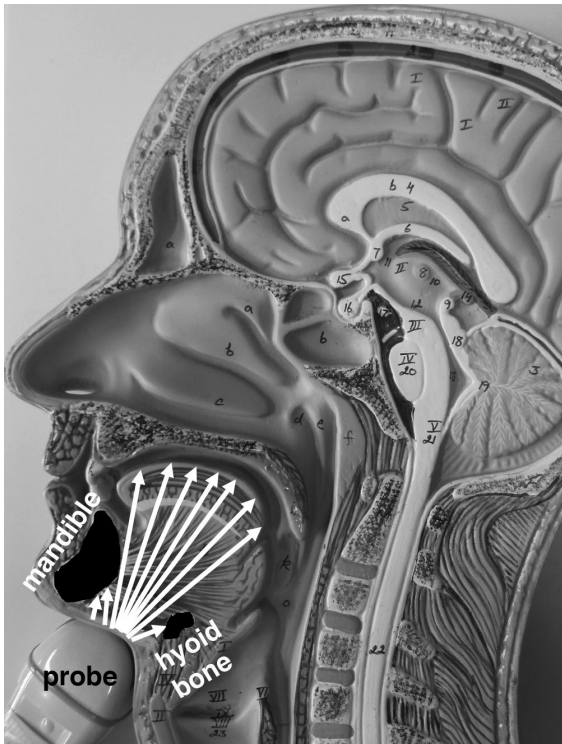


図-1 顎下に超音波プローブを当てた頭部正中矢状面画像。白い矢印は、超音波経路。黒色部は、超音波を吸収する下顎骨 (mandible bone) と舌骨 (hyoid bone)。

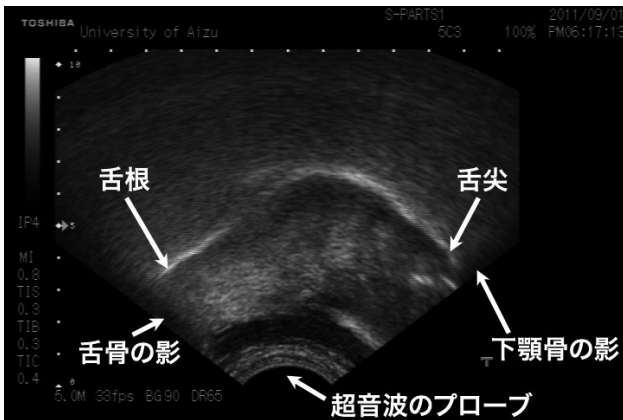


図-2 超音波による典型的な舌の正中矢状面画像。白い曲線は、舌の上の空気境界面を示す。舌尖は右側。

動いているように見えてしまう。それゆえ、正確性を期すためには、そのようなプローブの動きを制御、あるいは補正しなければならない。プローブと頭部の動きは、機械的制御により防ぐことができ、データ収集時に動きを追跡しておけば、後でソフトウェアでその動きを計算して取り除き、補正することもできる。今回のような解説記事では、頭部の相対位置の動きの制御と補正のための既存の方法について詳細には触れられないが、詳しくは [6-8] を参照されたい。超音波データ収集後の

分析方法については、定量的な舌形状分析のための方法が複数ある。幾つかの比較については [9] を参照されたい。

病院で使用される超音波機器は、通常、とても大きく、カートに載せられていて、高価である。音声研究室や言語聴覚士の診療室にはそのような機器を購入する資金も、設置スペースもないことが多い。しかしながら、幸い、超音波機器でも持ち運びが可能なポータブルモデルが数多く存在し、実験室や診療室の外でのフィールドワークに使用されている [10]。更には、USB 式の超音波プローブも存在していて、それらは、ノート型パソコンにつなぐだけで使用可能である。2013 年 11 月のウルトラフェスト (Ultrafest VI) では、Alan Wrench 氏が出席者に対して、現在どの超音波システムを使っているかのアンケート調査を行った。ウルトラフェストは、ほぼ隔年に開催されている国際会議で、特に超音波を使用している音声の研究者と教育者を対象としている。その会議で発表された Wrench 氏のアンケート調査の結果 [11] を、表-1 に示す。

発話中の舌の動きの超音波映像を記録するためには、超音波機器から出力される映像信号と、被験者が発話するマイクロホンからの音声信号を同期させる必要がある。使用されるハードウェアとソフトウェアによって、これら二つの信号の同期が問題となることが知られている。

映像信号については、機器内部のフレームレートは通常 100 フレーム/秒以上だが、機器に備え付けられた画面は標準的な映像の規格に制限され、その時間解像度は日本と北米では 30 フレーム/秒、ヨーロッパでは更に低く 24 フレーム/秒である。映像信号については、出力の種類や設定によっては、より高いフレームレートを得られることもある (例として、[12] を参照)。

超音波は、舌全体の観察を可能にするため、特に舌根の研究に適している。というのも、安全面、費用対効果面でも、他に例を見ない最適な方法であるからである。研究例の一つとしては [13] が挙げられるが、超音波を用いることで音韻素性と調音ジェスチャのマッピングの関係が明らかにされている。一度に舌全体を観察することで可能となる別の研究としては、調音器官同士のタイミングパターンの研究が挙げられる (例として [14] を参

表-1 世界中の音声研究室で使用される超音波機器の種類 (A. Wrench 氏のデータ [11] より)

タイプ	メーカー/モデル	使用研究室数
Cart-mounted (カート積載型)	Aloka SSD 1000/4000/5000/5500	4
	Zonare Z.one	1
	Toshiba Famio 8	1
Tabletop (テーブルトップ)	Mindray DP2200	6
	Mindray DP6600/6900	16
	GE Logiq alpha 100	1
Laptop (ノート型パソコン)	Mindray M5	1
	GE Logiq e	3
	Sonosite 180+	3
	Sonosite Titan	3
	Terason T3000	7
Tablet and probe (タブレット・プローブ)	Ultrasonix RP, Tablet, Touch	7
System plugs into laptop (ノート型パソコンプラグインシステム)	Echoblaster 128	3
Probe plugs into laptop (ノート型パソコンプラグインプローブ)	Interson SeeMore	2+

照)。このような研究では、舌尖、舌背、舌根など、舌の異なる部位における発話ジェスチャのタイミングの相対関係を究明しようとしている。

2. 超音波を用いた第二言語発音指導

今日、一般的な外国語の発音指導法及び学習法には、(a) 直感的・模倣的 (Intuitive-Imitative) アプローチ、及び (b) 分析的・言語学的 (Analytic-Linguistic) アプローチの 2 通りが存在する [15]。前者は、明示的な音声学的情報なしに、教師の声を模倣する伝統的なアプローチである。一方、後者は国際音声記号 (International Phonetic Alphabet, IPA)、調音表や調音情報、そして [16] にあるような声道模型などのツールを利用するアプローチである。最近の研究 [17] では、調音形状などの音声学的情報を明示的に与えることが、日本人学習者の英語の発音指導に有効であることが示されている。

コンピュータ支援言語学習 (Computer-Assisted Language Learning, CALL) のためのソフトウェアの普及につれ、音響信号は可視化され、発音モデルとの比較も可能となるだろう。しかしながら、音響と調音のマッピングは、しばしば解釈が難しく、学習者はどのように学習を進めて良いか分からずにいる。音響信号から学習者の調音を予測す

るようなシステムは存在するが、誤った予測をする可能性や、学習者の発音が誤っている際に正しい、あるいは、正しいときに誤っている、と判断してしまうような危険性がある (超音波を用いたそのようなシステムの検証に関する成功例は、[18] を参照されたい)。学習者の調音を直接計測し、バイオフィードバックを返すようなシステムは、そのような誤りが少ない。

音声学的知識を持たない言語学習者にとって、発話中の教師の正しい舌の動きを表示する映像は、スペクトログラムよりも、遥かに理解し易い。同じことは、画面上にリアルタイムで映し出される自分の舌の動きのバイオフィードバックについてもいえる。それゆえ、少なくとも 2006 年頃から、大学の発音指導や発音習得のクラスで超音波が取り入れられるようになった [19–21]。2009 年 1 月から 3 月まで (2010 年に再放送)、NHK 『3 か月トピック英会話』の番組では、会津大学の「語学研究センター音声学ラボ」(the CLR Phonetics Lab) が毎週レギュラーで取り上げられ、舌の超音波画像がテレビで放送されている。

日本人の学習者が、超音波画像を用いることで習得できると思われる英語の音は多数ある。例えば、日本語にない音である北米英語の /l/ と /r/ である。日本語のラ行子音は典型的に舌尖のみを使って発

音される。一方、英語の /l/ と /r/ は調音上、複雑な音である。/l/ は舌尖の上昇と舌背の後退という二つのジェスチャから成り、/r/ は舌尖又は舌端の上昇、舌根の後退、及び円唇化の三つのジェスチャから成る。学習者は、実際にこれら二つの英語の音の違い、また日本語の音との違いを捉えることで、それらの音を発音できるようになる可能性が高くなる。実際には、たとえ音の違いを聞き取れなくても、超音波フィードバックによって、各音を発音することはできるといふ仮説を立てる人もいるようである。日本語母語話者を対象とした超音波研究では、英語の /r/ と /l/ の聞き取りが、/r/ と /l/ の発音とは異なる過程で発達することが最近、明らかにされている [22]。

緊張/弛緩 (tense/lax) 母音の組み合わせにおいても、超音波は日本人英語学習者の役に立つであろう。例えば、/i/ と /ɪ/ (“heat”/“hit”), /u/ と /ʊ/ (“Luke”/“look”) などである。弛緩母音では、舌体がわずかに下がるだけでなく、緊張母音に見られるような舌根の前進はない。

プローブを 90° 回転させ、舌の冠状面を見ると、/s/ の発話時の舌の中心溝を見ることも可能であり、逆に /ʃ/ の発話時に舌の盛り上がりも観察できる。このようにして、学生は、例えば “seat” を “sheet” のように誤って発音した際に、どのように舌の形を直せば良いのか、見ることができる。

調音指導中に学習者に自分の舌を見せることは、超音波機器が 1 台しかない大きなクラスでは、当然、実施は困難である。しかしながら、舌のどこをどの程度動かすかなどの口頭での指示に、学習者が常に正しく従えるわけではないことを考慮すれば、リアルタイムにバイオフィードバックを与える利点は明らかである。超音波研究 [6] では、8 人の被験者に、ある一定方向に舌を動かしてもらう課題が行われた。被験者に対し超音波の画面を見ずに舌を動かす指示を与えた結果、何も見ないで指示どおりに舌を動かすことは非常に難しいことが分かった。しかしながら、画面に自らの舌が映し出されると、教師の指示に従って調音することが比較的スムーズにできた。将来的には、USB 型の超音波プローブの価格が下がり、大規模な発音クラスで多くのプローブが利用可能になることが期待される。

外国語発音指導における調音設定 (articulatory

setting), つまり、発話中の舌、唇、顎などにおいて各言語に潜む言語固有の設定 (setting) も、超音波の使用により研究が可能となる。[23] では、2 言語を母語話者並みに発音するバイリンガルにおいて超音波やその他のデータを調べた結果、各言語で異なる舌の安静位を持つことが明らかになった。この発見は、外国語の発音指導法において、将来的に重要な意味合いを持つ。

超音波は、聴覚障害者や言語治療を受ける人への発音指導にも役立つと考えられる。言語聴覚士の中には、患者の発話の診断と矯正のために、広範にわたり超音波を利用している人もいる (概説は [24], 又は [25–27] を参照)。

超音波を使用する音声研究者、臨床医、外国語発音指導者の数は過去十年で大きく増えた。超音波は、発話時の舌全体をリアルタイムで観察する安全で簡単な方法として、近い将来、価格の下降と共に、より多くの利用者を惹きつけていくであろう。

文 献

- [1] B. Gick, I. Wilson and D. Derrick, *Articulatory Phonetics* (Wiley-Blackwell, Chichester, UK, 2013).
- [2] M. Stone, “A guide to analysing tongue motion from ultrasound images,” *Clin. Linguist. Phonet.*, 19, 455–501 (2005).
- [3] C.A. Kelsey, F.D. Minifie and T.J. Hixon, “Applications of ultrasound in speech research,” *J. Speech Hear. Res.*, 12, 564–575 (1969).
- [4] 加藤正子, 大塚義顕, 向井美恵, 今富撰子, 木村智江, 出世富久子, “超音波前額断規格撮影法による側音化構音の観察,” *音声言語医学*, 43, 30–39 (2002).
- [5] N. Horiguchi and I. Wilson, “Head movement and tongue compression with a fixed ultrasound probe but a free head,” *Ultrafest 4*, New York University (2007).
- [6] N. Horiguchi, “How L2 pronunciation learners interpret articulation instructions: An ultrasound study of the tongue,” MSc Thesis, University of Aizu (2012).
- [7] T. Bressman, H. Flowers, W. Wong and J.C. Irish, “Coronal view ultrasound imaging of movement in different segments of the tongue during paced recital: Findings from four normal speakers and a speaker with partial glossectomy,” *Clin. Linguist. Phonet.*, 24, 589–601 (2010).
- [8] J. Mielke, A. Baker, D. Archangeli and S. Racy, “Palatron: A technique for aligning ultrasound images of the tongue and palate,” *Coyote Papers: Working Papers in Linguistics, Linguistic Theory at the University of Arizona*, 14, 96–107 (2005).
- [9] K.M. Dawson, D.H. Whalen and M.K. Tiede, “A comparison of methods for tongue shape analysis,” *Ultrafest 6*, University of Edinburgh (2013).
- [10] B. Gick, S. Bird and I. Wilson, “Techniques for field application of lingual ultrasound imaging,” *Clin. Linguist. Phonet.*, 19, 503–514 (2005).
- [11] A. Wrench, “Ultrasound speech analysis: State

- of the art,” *Ultrafest 6*, Edinburgh, Scotland (2013), http://materials.articulateinstruments.com/Technical/State_of_Art.ppt (参照 2014-07-12).
- [12] A.L. Miller and K.B. Finch, “Corrected high-frame rate anchored ultrasound with software alignment,” *J. Speech Lang. Hear. Res.*, 54, 471–486 (2011).
- [13] F. Hudu, “[ATR] feature involves a distinct tongue root articulation: Evidence from ultrasound imaging,” *Lingua*, 143, 36–51 (2014).
- [14] F. Campbell, B. Gick, I. Wilson and E. Vatikiotis-Bateson, “Spatial and temporal properties of gestures in North American English /r/,” *Lang. Speech*, 53, 49–69 (2010).
- [15] M. Celce-Murcia, D.M. Brinton, J.M. Goodwin and B. Griner, *Teaching Pronunciation*, 2nd ed. (Cambridge University Press, New York, 2010).
- [16] T. Arai, “Mechanical models of the human vocal tract,” *Acoust. Today*, 9, 25–30 (2013).
- [17] K. Saito, “Reexamining effects of form-focused instruction on L2 pronunciation development,” *Stud. Second Lang. Acq.*, 35, 1–29 (2013).
- [18] O. Engwall, “Analysis of and feedback on phonetic features in pronunciation training with a virtual teacher,” *Comput. Assist. Lang. Learn.*, 25, 37–64 (2012).
- [19] B. Gick, B.M. Bernhardt, P. Bacsfalvi and I. Wilson, “Ultrasound imaging applications in second language acquisition,” in *Phonology and Second Language Acquisition*, J.G. Hansen Edwards and M.L. Zampini, Eds. (John Benjamins, Amsterdam, 2008), pp. 309–322.
- [20] H.M.-L. Tsui, “Ultrasound speech training for Japanese adults learning English as a second language,” MSc Thesis, University of British Columbia (2012).
- [21] B. Meadows, “Do you see what I’m saying?: Ultrasound technology as a tool for pronunciation instruction,” in *Communication Technology for Students in Special Education and Gifted Programs*, J.E. Aitken, J.P. Fairley and J.K. Carlson, Eds. (Information Science Reference, Hershey, Pa, 2012), pp. 293–302.
- [22] M. Tateishi, “Effects of the use of ultrasound in production training on the perception of English /r/ and /l/ by native Japanese speakers,” MA Thesis, University of Calgary (2013).
- [23] I. Wilson and B. Gick, “Bilinguals use language-specific articulatory settings,” *J. Speech Lang. Hear. Res.*, 57, 361–373 (2014).
- [24] B. Bernhardt, B. Gick, P. Bacsfalvi and M. Adler-Bock, “Ultrasound in speech therapy with adolescents and adults,” *Clin. Linguist. Phonet.*, 19, 605–617 (2005).
- [25] M. Adler-Bock, B. Bernhardt, B. Gick and P. Bacsfalvi, “The use of ultrasound in remediation of English /r/ in two adolescents,” *Am. J. Speech-Lang. Pathol.*, 16, 128–139 (2007).
- [26] H.B. Klein, T. McAllister Byun, L. Davidson and M.I. Grigos, “A multidimensional investigation of children’s /r/ productions: Perceptual, ultrasound and acoustic measures,” *Am. J. Speech-Lang. Pathol.*, 22, 540–553 (2013).
- [27] J.L. Preston, N. Brick and N. Landi, “Ultrasound biofeedback treatment for persisting childhood apraxia of speech,” *Am. J. Speech-Lang. Pathol.*, 22, 627–643 (2013).



イアン ウィルソン (Ian Wilson)

会津大学語学研究センター教授。2006年、ブリティッシュコロンビア大学にて言語学博士号取得。2009年、2010年NHK「3か月トピック英会話」レギュラー。2013年、「調音音声学 (Articulatory Phonetics)」(Wiley-Blackwell Publishers 出版) 共著。大学院にて「音声の調音及び音響学 (Speech Articulation and Acoustics)」, 大学にて「発音」, 「論文指導」, 「英語 (English as a foreign language: EFL)」コースを担当。