ForumAcusticum
2 0 0 5

Articulation in the Didjeridu

Noam Amir
Department of Communications Disorders, Tel Aviv University, Tel Aviv, Israel, **noama@post.tau.ac.il**

The Didjeridu is well known for the wide range of timbres that it can produce. In order to produce these different timbres, beginning players are often instructed to shape the vocal tract as if articulating various vowel sounds. In reality, most novices find that actually producing a wide range of timbres is more difficult than what this simplified explanation would imply. Whereas beginners usually produce a rather "flat" tone, with practice their tone becomes clearer, and articulation improves. Expert players can produce a large range of clearly articulated timbres, with large variation in personal style. Good control of timbre involves subtle manipulations of the tongue, at its root, center and tip, the mouth cavity as controlled by the cheeks, and the lip shape and tension. Such control can take extensive practice to master, and accurate measurements of the physical parameters involved is probably very difficult.

On the other hand, the acoustic manifestations of such accurate control in the resultant spectrum are far easier to evaluate and measure. Despite this, current literature has not examined this in detail, beyond pointing out formant-like peaks in the spectral envelope.

In the present study we present analyses carried out on sustained notes of several expert didjeridu players, examining the spectral parameters which characterize the highly articulated timbre that these players can produce.

1 Introduction

The Australian Didjeridu (or Didgeridoo) has gained a large degree of popularity in recent years. Its appeal is probably due to the fact that its tone is very different in character than most western instruments. Despite the fact that it plays mainly one tone, a proficient player can produce a very large range of sounds from this instrument. These range from percussive sounds, created by blowing various types of bursts into the instrument, to long meditative tones. The latter, steady state tones can be produced with a very large variability in timbre, which is quite rare in western instruments.

It has been clarified by now [1, 2, 3] that the large variability in timbre is due mainly to the strong coupling between the instrument and the player's vocal tract. The throat end of the didjeridu is much wider than that of any western brass instrument, and also lacks a mouthpiece. This brings the characteristic input impedance at the throat to a much lower level than that found in these brass instruments. This greatly increases the influence of the impedance of the vocal tract on the resulting sound.

Students new to the Didjeridu are often instructed to shape their vocal tract as if they are articulating different vowel sounds, in order to manipulate the timbre of the sound. From personal experience, accumulated through several years of playing, there are two difficulties associated with actually carrying this out. The first, more minor difficulty is that the didjeridu sound as perceived by the player can be quite different from the sound heard by other listeners, due to peculiarities in the internal conduction of sound in the head, vs. externally radiated sound. The second, major difficulty is that actually achieving the distinct timbres is far more difficult than simply shaping the vocal tract as in speech. Many subtle adjustments are involved in placement of the tongue and tensioning and placement of the lips. Students often complain that one of the major difficulties in attaining proficiency on the Didjeridu is that of getting good control over the instrument's timbre.

Measuring the subtle adjustments made by an advanced player is clearly extremely complicated, if not impossible, though some studies are and have been carried out in this respect, on the didjeridu and brass wind instruments, through photographing the lips by various means or replacing them with artificial blowing devices [3]. On the other hand, careful and detailed analysis of the sounds that can be produced on the Didjeridu has only begun to be carried out.

Previous studies have described formant like peaks in the spectral envelope at moderately high frequencies, as the result of changes in shape of the player's vocal tract [2,3]. After several years of playing the Didjeridu, I have gradually come to the conclusion that there are many subtleties in the didjeridu timbre that have not been examined in detail. This has been examined with respect to low frequencies in a recent paper [5]. On the other hand, it is the high frequencies that contribute to a distinct sense of "sharpness" or distinctness in the timbres that good players can produce. In such cases there is sometimes a sense of similarity to vowel sounds such as /a/, /i/, /u/, etc., which is probably the aspect of timbre attributable to the existence of formants.

The purpose of this initial study is therefore to examine formant phenomena present in the higher spectrum of the didjeridu. We wish to observe how these formants relate to the perceived sound, and obtain an idea as to the complexity or dimensionality of this phenomenon. One could envision that it would be possible to find a simplified measure for the "sharpness" of the spectrum that can be used to present a beginning player with visual feedback indicating the quality of his or her tone.

2 Material

Three players were asked to play a range of timbres on a single instrument. The players were Y, a professional player and teacher (the instrument was his own personal Didjeridu, in the key of D, which he is very accustomed to and probably plays more than any other instrument), R, a very accomplished player, and T, also an excellent player and performer. All three were requested to play as wide a range of timbres as they can. It is interesting to note that Y and R felt comfortable and produced such a range easily, whereas T complained that he was unaccustomed to the instrument and had been playing an instrument in F a great deal lately. Though he had no problem in producing a clean tone, he had difficulties in producing distinct timbres.

All players were recorded to the disk of a laptop computer, using a JVC MV-40 (cardioid) microphone and an external soundcard (Soundblaster Audigy), at a sampling rate of 48 kHz, 16 bits per sample.

3 Analyses

3.1 Spectra

All players produced pitches of 73 ± 0.5 Hz. A sliding analysis window of 100ms was applied (approximately 7.5 periods), and multiplied by a Blackman weighting function. An FFT of length 65536 was then applied to each window, giving spectral resolution of about 1/8 Hz. A Matlab program was written by the author, in order to enable convenient scrolling through the signal and viewing of local spectra. Due to the fact that very little spectral information was present above 8 kHz, all the recordings were downsampled to 16kHz.

Player Y was asked to produce a series of tones, from a very "flat" tone through steadily stronger and perceptually higher (in frequency) articulations, up to a quite distinct /i/ tone. The spectra of these tones are presented in Figure 1, from top to bottom. We observe several interesting phenomena:

1. The flat tone has a spectrum that decreases nearly uniformly in amplitude from the 8'th harmonic and beyond.
2. Tones 2 to 6 have two more or less distinct formants, that seem to get progressively closer and closer to each other in frequency, with the first formant rising in frequency and the second one descending. Interestingly, in normal speech, when vocalizing a progression from /a/ to /i/ the *second* formant follows a trajectory that is vaguely similar to that of the Didjeridu's *first* formant.
3. The two formants do not seem to be able to merge, so that at the terminal /i/ tone there is still a sharp dip in amplitude between them.
4. At low frequencies the behavior is nearly uniform across all tones. The fundamental and 4^{th} harmonic are strong, whereas nearly all the other harmonics are relatively weak, up to the so called formant. This has been discussed in detail in a previous study [5]

To summarize, the spectra are more complicated in shape than what would be implied by the presence of a single formant. A more detailed perceptual analysis of the importance of the various features in the spectrum is beyond the scope of this current paper, though it would be interesting to examine this in the future.

The results mentioned above imply that it could be relatively straightforward to define a simplified measure of the degree to which the perceived sound has a clear articulation. In order to facilitate the latter we turned to a parametric description of the spectral envelope.

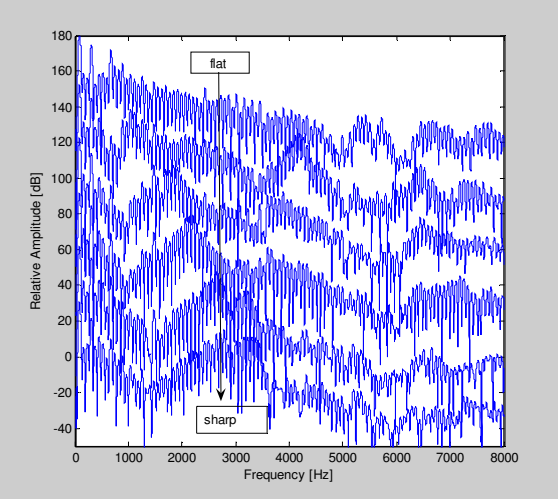


Figure 1: Spectra of successive tones for player Y, from flat (top trace) to sharp /i/ tone (bottom trace)

3.2 Modeling the spectral envelope

To examine these points further, a 10th order all-pole (LPC – Linear Predictive Coding) estimate of the spectral envelope was calculated. This model is especially suited to the modeling of spectral peaks. The spectral envelopes calculated on the basis of these models for the same 6 tones are shown in Figure 2. The formant-like phenomena outlined above are more easily observed here.

It should be pointed out though, that this method is actually a rather poor one for resynthesis of the didjeridu tone. It gives a grossly oversimplified spectral envelope, especially at low frequencies, where strong harmonics are adjacent to weak ones.

In any case, it seems that the two main formants between 1000 and 5000 Hz can be used to give an indication to the player regarding the degree of articulation which he has achieved. This can serve as a form of visual feedback for beginning players who encounter difficulties in this respect. Poles that correspond to strong spectral peaks are located in the z-plane at points very near to the unit circle. This can be utilized to give an indication of the formant center frequencies and peakiness.

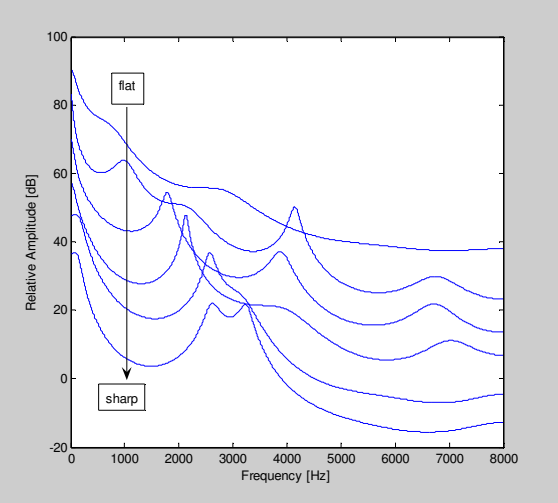


Figure 2: Spectral envelopes of successive tones for player Y, from flat (top trace) to sharp /i/ tone (bottom trace)

3.3 The "Articulation Index"

It is well known that poles in the all-pole model that correspond to strong spectral peaks are located in the z-plane at points very near to the unit circle. Clearly, the "flat tone" will have no such dominant poles, whereas all the other spectra have dominant poles at angles corresponding to the range between 1000 and 5000 Hz. This is demonstrated in Figure 3. Therefore, the distance between these poles and the unit circle can serve as an "articulation index" which will give the player visual feedback as described above. A good measure should be based on the reciprocal of the distance of the poles from the unit circle, since the spectral peaks are roughly proportional to this value.

The measure we propose here is simply the product of the above reciprocals for the two dominant poles in the 1000 to 5000 Hz frequency band. In other words, for two complex poles P_1 and P_2, the articulation index A is:

$$A = \frac{1}{1-|P_1|} \cdot \frac{1}{1-|P_2|} \qquad (1)$$

The articulation index over a succession of tones played by the three players (Y, R, T) is presented in Figure 4, together with the separate reciprocals for the two poles. The index for the first part of player Y's excerpt is very low, corresponding to a "flat tone". For the rest of his excerpt it is mostly in the range 200 to 400. Player R, which is less experienced and less accustomed to the instrument he played on, has a lower index, with values in the vicinity of 100 on his steady state tones. Player T, on the other hand, which complained of being unaccustomed to the instrument, has a consistently low index of around 20.

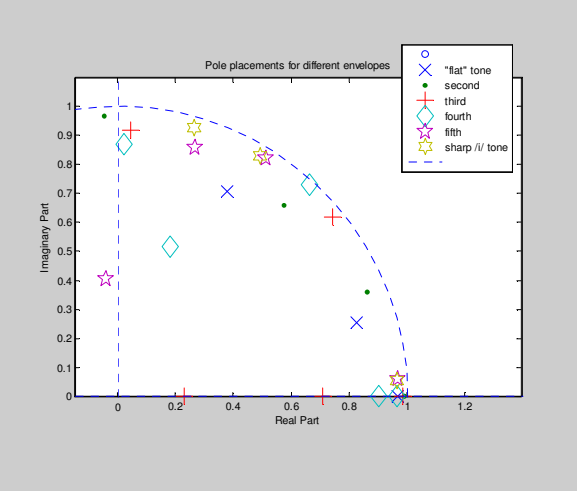


Figure 3: Poles in first quadrant for allpole model of successive tones for player Y, from flat to sharp /i/ tone

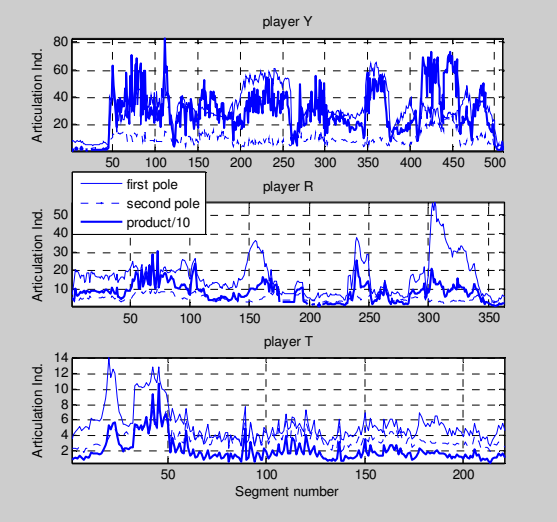


Figure 4: Articulation indices for excerpts taken from 3 players, Y, R and T.

4 Conclusions & further research

The neutral or "flat" timbre of the didjeridu decays more or less uniformly in the spectrum, beyond approximately the 9th harmonic. Strongly articulated tones, on the other hand, have concentrations of energy in formant-like regions. The details of the spectrum can be rather complicated; for example, as we observed in the playing of one expert player, raising the first formant lowers the second formant, until they coincide, though they do not merge. The *perceptual* importance of the spectral details, on the other hand has not been yet evaluated fully. In the recordings examined here, the dominant poles in the allpole model of the spectral envelope, as applied usually to speech signals, can be used to give a good indication of the degree to which the articulation of the sound is clearly defined.

Though we have shown the overall behavior of the spectrum for a small number of players on a single instrument, this is far from an exhaustive study. It will be very interesting to examine whether the behavior of the two formants found above, generalizes to other cases, such as: other players on the same instrument, other instruments in the same key, and instruments in other keys. The variation in tone across players and instruments is quite astounding at times, and it remains to be observed what is common across all players, or good players vs. poor ones. The perceptual importance of each separate formant is also a point to consider.

The physical mechanisms for the behavior of the above two formants should be examined in further work. In speech, the first two formants serve to identify a vowel sound, whereas it seems that in the Didjeridu the perceived semi-vowel sound is based solely on the first formant, corresponding to the second speech formant. It is interesting to examine whether the player can control the second Didjeridu formant, corresponding in frequency range to the third speech formant.

There is a certain dual character to the Didjeridu's first formant. Sometimes it is perceived almost as a distinct pitch, as in overtone singing, and sometimes as a distinct timbre giving it vowel type characteristics. Correlating these perceptual characteristics to the acoustic ones is also a subject that should be examined further.

Acknowledgments

I'd like to thank Yoram Sivan, Rotem Koren and Tal for their patience and willingness to participate in this study.

References

[1] N. H. Fletcher, “Acoustics of the Australian didjeridu”, *Australian Aboriginal Studies*, Vol. 1, pp 28-37, 1983

[2] N. H. Fletcher, “The didjeridu (didgeridoo)”, *Acoustics Australia*, Vol. 24, pp. 11-15 (1996)

[3] N. H. Fletcher, L. Hollenberg, J. Smith, J. Wolfe, “The Didjeridu and the vocal tract”, *Proceedings of ISMA 2001*, Perugia (2001)

[4] Amir, N., Alon, Y., “A study of the didjeridu: normal modes and playing frequencies”, *Proceedings of ISMA 2001*, Perugia (2001)

[5] Amir, N. "Some insights into the acoustics of the Didjeridu" *Applied Acoustics*, Vol. 65, pp 1181-1196 (2004)