

BAB 5

PENUTUP

Bab ini terdiri dari dua bagian. Bagian 5.1 berisi kesimpulan hasil penelitian penulis selama pelaksanaan tugas akhir. Untuk perkembangan penelitian SPS Bahasa Indonesia di masa mendatang, penulis juga menuliskan beberapa saran pada bagian 5.2 yang juga merupakan bagian akhir dari bab ini.

5.1 Kesimpulan

Pada penelitian ini, penulis melakukan dua jenis eksperimen. Pertama, penulis melakukan eksperimen menggunakan enam kamus fonetik, yaitu IPA, SAMPA, ARPABET, kamus Lestari, kamus Sakti, dan kamus Zahra, dalam rangka memperoleh kamus fonetik yang paling tepat untuk SPS Bahasa Indonesia. Kedua, penulis melakukan eksperimen melibatkan data rekaman berita RRI. Dari hasil eksperimen dan analisis yang telah dilakukan, penulis menarik sejumlah kesimpulan, antara lain:

1. Kamus fonetik yang paling tepat untuk merepresentasikan bunyi pada SPS Bahasa Indonesia adalah kamus Zahra, yang berisi 33 simbol, terdiri dari 6 vokal, 4 diftong, dan 23 konsonan, dengan beberapa catatan sebagai berikut:
 - Bunyi huruf “e” cukup direpresentasikan dengan dua alofon.
 - Bunyi huruf “o” cukup direpresentasikan dengan satu alofon.
 - Adanya tambahan alofon /ey/ untuk melengkapi bunyi diftong dalam kamus fonetik.
 - Adanya tambahan alofon /sy/, /kh/, /ny/, dan /kk/ untuk melengkapi bunyi konsonan dalam kamus fonetik.
 - Alofon /v/ tidak perlu digunakan, cukup alofon /f/ untuk merepresentasikan bunyi huruf “v” dan “f”.
2. Dengan menggunakan 1.000 berkas rekaman telepon seluler dalam proses pelatihan, pengujian terhadap 250 berkas rekaman telepon seluler

menggunakan kamus Zahra menghasilkan akurasi maksimum yang dapat diperoleh sebesar 73,5%. Jumlah kalimat yang dikenali dengan benar sebanyak 138 kalimat.

3. Pengujian terhadap 100 berkas rekaman berita RRI, dengan OOV (*Out of Vocabulary*) sebesar 8,92%, menghasilkan akurasi sebagai berikut:
 - 71,1% untuk pengujian yang didahului pelatihan menggunakan 1.000 berkas rekaman telepon seluler. Jumlah kalimat yang dikenali dengan benar sebanyak 33 kalimat.
 - 71,6% untuk pengujian yang didahului pelatihan menggunakan 1.010 berkas suara, yaitu 1.000 berkas rekaman telepon seluler dan 10 berkas rekaman berita RRI. Jumlah kalimat yang dikenali dengan benar sebanyak 33 kalimat.

Dengan demikian, terjadi peningkatan akurasi sebesar 0,5% pada pengujian terhadap 100 berkas rekaman berita RRI

5.2 Saran

Untuk pengembangan Sistem Pengenalan Suara (SPS) Bahasa Indonesia yang lebih baik di masa mendatang, penulis memiliki beberapa saran sebagai berikut:

1. Perekaman data suara dilakukan dalam ruangan hening, sehingga tidak ada lagi gangguan berupa suara obrolan orang lain, dering ponsel, batuk, suara kendaraan, suara tawa, dan lain-lain yang muncul bersamaan dengan suara utama dari pembicara.
2. Gunakan data suara dengan *sample size* 8 bit dan *sample rate* 8 kHz dan data suara dengan *sample size* 16 bit dan *sample rate* 16 kHz untuk pelatihan, sehingga SPS dapat memperhitungkan data suara dalam berbagai kondisi.
3. Tambah jumlah data suara untuk pelatihan, sehingga seluruh trيفون Bahasa Indonesia dapat tercakup dalam model akustik hasil pelatihan. Dengan demikian, tidak hanya monofon saja yang lengkap tercakup di dalam

korpus, melainkan juga trifon, sehingga akurasi SPS meningkat. Untuk beberapa aplikasi SPS Bahasa Inggris yang bersifat komersil bahkan telah mencakup bunyi hingga level pentafon (*quintphones*).

4. Transkripsi harus dipastikan sama dengan isi data suara. Hindari kesalahan penulisan transkripsi semaksimal mungkin.
5. Hindari penggunaan kata-kata yang asing menurut ejaan Bahasa Indonesia, baik dalam korpus teks untuk model bahasa maupun korpus suara, misalnya “aap”, “mujhe”, dan lain-lain. Walaupun tetap ada kata-kata seperti ini, buatlah pelafalannya secara manual dengan mendengarkannya satu per satu, sehingga cocok dengan pengucapan oleh pembicara. Akan lebih baik jika kata-kata yang terkandung dalam korpus teks dan suara adalah kata-kata yang sesuai dengan Ejaan Yang Disempurnakan ditambah kata-kata pada percakapan sehari-hari.
6. Lakukan penyeragaman kata, sehingga tidak ada lagi kata-kata berbeda namun berlafal dan bermakna sama. Sebagai contoh, “triji”, “3G”, “three ji”, “three g” diseragamkan menjadi “3G”.
7. Perbesar ukuran korpus teks untuk pembuatan model bahasa, sehingga kosakata yang dimiliki semakin kaya. Penambahan data teks dapat dilakukan dengan mengambil artikel-artikel koran dan majalah. Konsekuensi dari pembesaran ukuran korpus teks adalah penambahan kata-kata beserta lafalnya ke dalam kamus pelafalan, dan hal ini tidak dapat dilakukan secara otomatis seluruhnya. Oleh karena itu, penulis menyarankan untuk merangkul sejumlah tenaga untuk membuatnya, tidak dari nol, melainkan memeriksa dan memperbaiki pelafalan kata hasil pembuatan kamus pelafalan dari suatu program sederhana berdasarkan kamus fonetik pada Tabel 4.11.
8. Perkaya dialek pembicara yang direkam menjadi korpus data suara dengan dialek dari berbagai suku bangsa di Indonesia, seperti Batak, Sunda, Jawa, Madura, Makassar, dan sebagainya. Dengan demikian, SPS dapat mengenali ucapan seluruh warga Indonesia, tidak hanya ucapan warga Jakarta. Walaupun korpus data suara yang digunakan pada pelaksanaan

tugas akhir ini berasal dari pembicara dengan variasi suku bangsa, namun dialek daerahnya sebagian besar sudah hilang.

