Predicted Exam Answers and Q

1. Semantic Web

Q1: Explain the difference between RDF and RDFS with examples.

- **Answer:**
    - **RDF (Resource Description Framework):**
      Focuses on representing data using triples. Example:
      (John, hasPet, Dog) — represents that John has a pet dog.
    - **RDFS (RDF Schema):**
      Adds semantics by defining classes and properties. Example:
        - <John> belongs to class <Person>.
        - <Dog> belongs to class <Animal>.
        - Property <hasPet> links <Person> to <Animal>.

Q2: Why is URI important in ensuring uniqueness in the Semantic Web?

- **Answer:**
  URIs act as unique identifiers for resources. Without them, entities like <John> in one dataset could conflict with <John> in another. URIs ensure data integration by maintaining global uniqueness.

Q3: Create an RDF graph for the statement:

- "Alice likes chocolate, and chocolate is a type of food."
- **Answer:**
    - Triples:
        1. <Alice> <likes> <Chocolate>
        2. <Chocolate> <type> <Food>

---

2. Source-Filter Model

Q1: Compare the roles of the vocal cords (source) and the vocal tract (filter) in producing voiced and voiceless sounds.

- **Answer:**
    - **Voiced Sounds:**
        - Source: Vocal cords vibrate to produce sound waves.
        - Filter: Vocal tract modifies these waves into phonemes like vowels (/a/, /e/).
    - **Voiceless Sounds:**
        - Source: No vocal cord vibration (e.g., /s/, /f/).

        - Filter: Vocal tract still shapes airflow into specific sounds.

Q2: What happens to the speech signal when the filter (vocal tract) shape changes rapidly?

- **Answer:**
  Rapid changes in the vocal tract alter formant frequencies, creating transitions that distinguish diphthongs (e.g., /ai/ in "kite") or rapid consonant-vowel sequences.

---

3. Text-to-Speech

Q1: What are the challenges in concatenating speech units in unit selection TTS systems?

- **Answer:**
    - Discontinuity between units: Abrupt changes in pitch or energy.
    - Lack of smoothness: Selected units may not blend naturally.
    - Limited database coverage: Some phonetic contexts may not exist in the database.

Q2: How is prosody (intonation, stress) incorporated into modern TTS systems like WaveNet?

- **Answer:**
    - Prosody is modeled by analyzing linguistic context (e.g., sentence structure) and embedding pitch contours or duration patterns directly into neural networks.
    - WaveNet generates audio samples frame-by-frame, ensuring natural intonation and stress.

---

4. Multimodal Systems

Q1: In a smart home, describe how a multimodal system can prioritize visual (gesture) and auditory (voice) inputs in a noisy environment.

- **Answer:**
    - **Scenario:** In a noisy room, voice commands might be unclear.
    - **Solution:**
        - The system detects high ambient noise using audio sensors.
        - Prioritizes gestures (e.g., waving to turn off lights) over voice commands.

Q2: How does adding emotion recognition improve the effectiveness of a multimodal system?

- **Answer:**
    - Emotion recognition adds context to user inputs.

- Example: If a user's tone is frustrated, the system might prioritize speed and brevity in its response.

---

5. Acoustic Phonetics

Q1: Given a spectrogram, identify the difference between a stop sound and a fricative.

- **Answer:**
  - **Stop Sound:**
    - A short burst of noise (e.g., /p/, /t/).
    - Appears as a sudden spike in energy followed by silence.
  - **Fricative:**
    - Continuous turbulence (e.g., /s/, /f/).
    - Appears as consistent high-frequency energy over time.

Q2: Explain what formant transitions indicate about the movement of the tongue during speech.

- **Answer:**
  - Formant transitions show how the tongue moves between vowel and consonant positions.
  - Example: In /ba/, the transition from the /b/ closure to the vowel indicates the tongue rising.

---

6. General Questions

Q1: How can advances in Semantic Web technologies improve voice-based search engines?

- **Answer:**
  - Semantic Web allows search engines to understand context and relationships in data.
  - Example: Instead of returning links, a query like "Who is the CEO of Tesla?" can directly return "Elon Musk" by analyzing RDF data.

Q2: Propose a new application combining Text-to-Speech, Multimodal Systems, and Acoustic Phonetics for inclusive education.

- **Answer:**
  - **Application:** A smart assistant for hearing-impaired students.
    - **Text-to-Speech:** Converts typed text into audible speech for other students.

- **Multimodal:** Allows input via gestures or text.
- **Acoustic Phonetics:** Analyzes speech clarity for pronunciation practice.