The diagram compares the DER (%) for different systems under three conditions: Clean speech, Speech + noise, and Speech + music. The x-axis represents the systems (System 1 to System 5), and the y-axis represents the DER (%) ranging from 0 to 60.

- **System 1**: Shows a moderate DER (%) for all conditions, with the highest DER (%) in the Speech + noise condition.
- **System 2**: Demonstrates the highest DER (%) in the Speech + noise condition, significantly higher than in other conditions.
- **System 3**: Exhibits a relatively lower DER (%) across all conditions, with a slight increase in Speech + music.
- **System 4**: Displays a balanced DER (%) across conditions, with slight increases in Speech + noise and Speech + music.
- **System 5**: Shows a moderate DER (%) in all conditions, with a slight rise in Speech + music.

Overall, System 2 performs the poorest under the Speech + noise condition, whereas System 1 has a more consistent performance across all conditions.