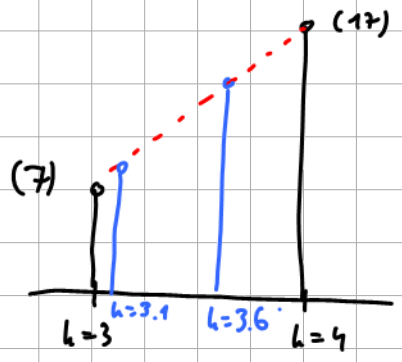
1)



Linear interpolation

$$X(h) = m \cdot h + n$$

$$m = \frac{17-7}{4-3} = 10$$

$$7 = 10 \cdot 3 + n \iff n = 7 - 30 = -23$$

$$\Rightarrow X(h) = 10 \cdot h - 23$$

$$X(h=3.1) = 10 \cdot 3.1 - 23 = 31 - 23 = 8$$

$$X(h=3.6) = 10 \cdot 3.6 - 23 = 36 - 23 = 13$$

3) One Bandpass per Bark ⇒ 24 Bandpasses

10 ms hop size ⇒ 100 Spectra per second

⇒ Datarate = $100 \frac{\text{Spectra}}{\text{s}} \cdot 24 \frac{\text{Bandpasses}}{\text{Spectrum}} \cdot 4 \frac{\text{Byte}}{\text{Bandpass}}$

$= 100 \cdot 24 \cdot 4 \cdot \frac{\text{Byte}}{\text{s}} = 9600 \frac{\text{Byte}}{\text{s}}$

Audio CD: $2 \cdot 44100 \cdot 16 \frac{\text{Bit}}{\text{s}} \cdot \frac{1}{8} \frac{\text{Byte}}{\text{Bit}} = 176\,400 \frac{\text{Byte}}{\text{s}}$

MP3 roughly factor 10 more compact than Audio CD:

$17640 \frac{\text{Byte}}{\text{s}}$

Voice Codec for GSM: roughly $12\,000 \frac{\text{Bit}}{\text{s}}$

$= 1500 \frac{\text{Byte}}{\text{s}}$

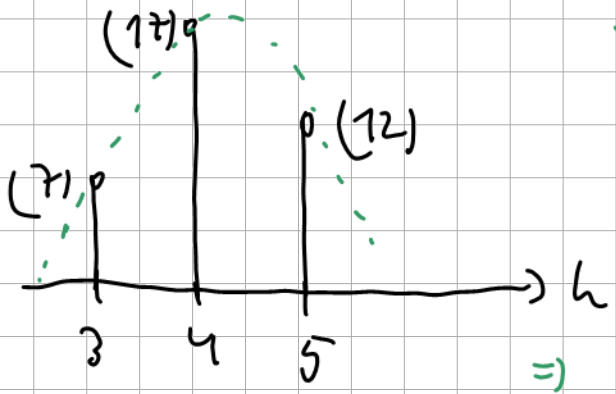
4) Yes, for a wide range of time stretching human voice stays understandable.
Human speakers also use different speed when they are talking.

5) Yes, human speech stay understandable for a wide range of pitches.

Humans are able to understand children (which usually has the highest fequencies) and people with very low pitches, like bass-singers.

6) Transmitting only the envelope of local spectra results in very low quality but human speech stays understandable (in general).

2)



$X_{(h)} = a\,h^2 + b\,h + c$

$7 = a \cdot 9 + b \cdot 3 + c$
$17 = a \cdot 16 + b \cdot 4 + c$
$12 = a \cdot 25 + b \cdot 5 + c$

$\Rightarrow \quad 10 = 7a + b \qquad \cdot (-2)$
$\Rightarrow \quad 5 = 16a + 2b$

$\Rightarrow \quad -15 = 2a$
$\Rightarrow \quad a = -7{,}5$

$b = 10 - 7 \cdot (-7{,}5) = 62.5$

$c = 7 - 9 \cdot (-7{,}5) - 3 \cdot 62.5 = -113$

$X_{(h=3.1)} = -7.5 \cdot 3.1^2 + 62.5 \cdot 3.1 - 113 = 8.68$

$X_{(h=3.6)} = -7.5 \cdot 3.6^2 + 62.5 \cdot 3.1 - 113 = 14.8$

7) Pitch is irrelevant, human voice stays understandable, even if the pitch is modified.
The speed of talking is irrelevant: human voice stays understandable for slow speakers and fast speakers in a wide range of talking speed.
The envelope of the local spectrum is relevant for voice understanding, as shown by reducing the number of bandpasses per bark to a very small value.