The role of anticipatory and perseveratory coarticulation cues for speech perception in young and old listeners

Rachel Haines[1], Heather Fortnum[1] and Antje Heinrich[2] .

[1]Nottingham Hearing Biomedical Research Unit, UK. [2]MRC Institute of Hearing Research, Nottingham, UK.

NHS National Institute for Health Research

This work was funded by an NIHR PhD studentship to Rachel Haines.

MRC Institute of Hearing Research

INTRODUCTION

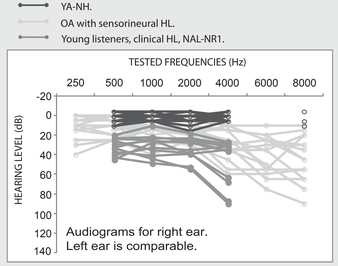
- Coarticulation is the partial spectral-temporal overlap of speech segments.
- Coarticulatory cues benefit speech perception of normal-hearing young adults in silence (Ostreicher & Sharf, 1976) and potentially in noise (Bronkhorst et al. 1993).
- Coarticulation can be differentiated into anticipatory (ANT) and perseveratory (PERS) cues.
- A dissociation of ANT and PERS cues in patients with motor speech disorder (dysarthria), who show preserved production of ANT cues but disrupted production of PERS cues, suggests that these cues might be governed by different processes.
- Based on production, it has been suggested that ANT cues rely on higher-level phonetic processing while PERS cues rely on biomechanical motor processes and constraints (McCaffrey Morrision, 2008 , Baum & Waldstein, 1991).

This study investigates:
1) whether listeners benefit perceptually from ANT and PERS cues in quiet and in noise.
2) whether the dissociation found for ANT and PERS cues in production holds for perception, and how the pattern of benefit differs between these cues.
3) how perception of ANT and PERS cues is affected by age, hearing status and noise.

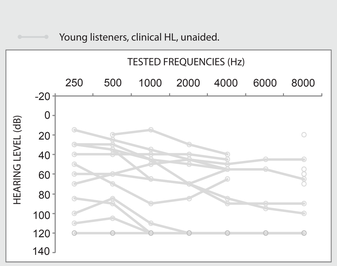
METHODS

Participants

- 16 young adults with normal hearing (YA-NH); mean age: 25.7 years of age.
- 18 older adults with age-normal hearing (OA); mean age: 64.9 years of age.
- 12 young adults with clinically significant hearing impairment (YA-HI); mean age: 28.4 years of age. Their listening was aided.



Audiogram for YA-NH and OA. Aided YA-HI calculated from NAL-NR1.



Unaided audiogram for YA-HI.

METHODS cont'd

Stimuli

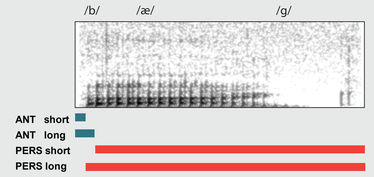
- 36 C_1VC_2 words starting in **ANT** with: /b, d/ (voiced) or /t, s/ (unvoiced).
  in **PERS** with: /b, d/ (voiced) or /t, k, s, sh/ (unvoiced).
- Stimuli shortened to only C_1 (ANT) or VC_2 (PERS).
- Two coarticulation cue durations used: shorter (1 periodic striation) versus longer (2 periodic striations).
- 8-talker babble with a 2s lead-in, 1s end and 10ms rise- and –fall time.
- All stimuli were presented 65dB SPL, either in quiet or at +5dB above individually-determined 50% SRT threshold.

Task

- Two words displayed on the screen, followed by auditory presentation of edited stimuli C_1 or VC_2.
- Participants decided which word represented the un-cut version of the stimulus.

ANT Trial Example
Target Stimulus: 'b'
Word Choices: 'bag', 'boot'

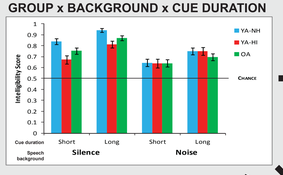
PERS Trial Example
Target Stimulus: 'ag'
Word Choices: 'tag', 'bag'



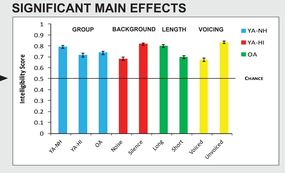
RESULTS

Anticipatory Coarticulation (ANT)

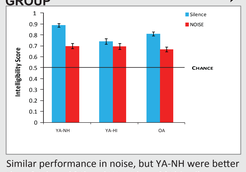
GROUP x BACKGROUND x CUE DURATION



SIGNIFICANT MAIN EFFECTS

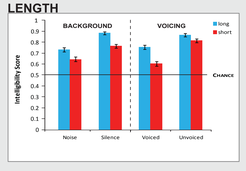


SIGNIFICANT INTERACTIONS WITH GROUP



Similar performance in noise, but YA-NH were better than OA ($p < .001$) and YA-HI ($p < .001$) in silence.

SIGNIFICANT INTERACTIONS WITH LENGTH



Overall Analysis

Factor	Hearing sensitivity as covariate NO	YES
Group	(F(2,8450)=6.7**	n/s
BKG	(F(1,8450)=215.79***	(F(1,8450)=75.31***
LEN	(F(1,8450)=115.81**	(F(1,8450)=34.03***
VOI	(F(1,8450)=271.55***	(F(1,8450)=106.8***
GROUP*BKG	F(2,8450)=28.12***	F(2,8450)=4.59**
GROUP*VOI	n/s	F(2,8450)=3.5*
BKG*LEN	(F(1,8450)=15.65***	(F(1,8450)=17.07***
VOI*LEN	(F(1,8450)=8.63**	(F(1,8450)=8.5**

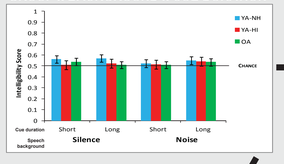
	p<
*	0.05
**	0.01
***	0.000

Analysis per Group

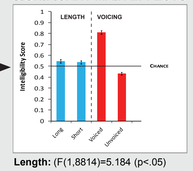
	YA-NH vs OA		YA-NH vs YA-HI		YA-HI vs OA	
	Covariate NO	YES	Covariate NO	YES	Covariate NO	YES
Group	**	n/s	**	n/s	n/s	n/s
BKG	***	***	***	**	**	***
LEN	***	***	***	**	**	***
VOI	***	***	***	**	*	***
GROUP*BKG	***	n/s	***	***	**	n/s
GROUP*LEN	*	*	n/s	n/s	n/s	n/s
GROUP*VOI	n/s	n/s	n/s	**	n/s	n/s
BKG*LEN	***	***	**	**	**	***
VOI*LEN	**	**	*	*	n/s	n/s

Perseveratory Coarticulation (PERS)

GROUP x BACKGROUND x CUE DURATION

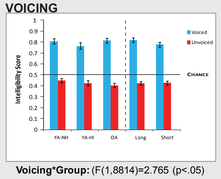


SIGNIFICANT MAIN EFFECTS



Length: (F(1,8814)=5.184 (p<.05)
Voicing: (F(1,8450)=905.468 (p<.001)

SIGNIFICANT INTERACTIONS WITH VOICING



Voicing*Group: (F(1,8814)=2.765 (p<.05)
Voicing*Length: (F(1,8814)=6.232 (p<.05)

Using hearing sensitivity as covariate did not change the pattern of results for PERS.

Data Analysis

Binary linear regression with fixed factors **Group** (YA-NH, OA, YA-HI), **Background** (Silence, Babble noise), **Coarticulation cue length** (Short, Long), and **Voicing of C_1** (Voiced, Unvoiced), separately for ANT and PERS.

Hearing sensitivity (PTA for YA-NH & OA; calculated NAL-NR1 gain prescriptions for YA-HI) was entered as a covariate in a second step.

DISCUSSION

1) Listeners can use both ANT and PERS coarticulation cues for speech perception.
2) All experimental variables affected ANT cues, but only voicing affected PERS.
   This can be explained by the nature of the acoustics: At release of C_1, unvoiced segments contain a much stronger and longer release transient compared to voices segments, whose main acoustic informastion is located at the onset of the vowel transition, which occurs later in the sound. Cue duration will particularly affect perception of voiced stimuli in ANT conditions because perception relies on the preservation of the vowel transition. In the shorter stimulus length in ANT trials this information is not preserved. It also explains why performance is at chance for unvoiced PERS trials: the crucial information that would have let listeners differentiate between alternatives is located at the onset of the stimulus and was removed in PERS trials.
3) Audibility and age have different effects on ANT and PERS cues.
   ANT: performance was poorer for OA and YA-HI in several conditions compared to YA-NH. The **overall group** differences are due to audibility. The **group x background** interaction is reduced when audibility is taken into account, and indeed the age effect between YA-NH and OA is due to it. YA-HI's use of coarticulation cues remains reduced compared to that of NH listeners of all ages even when their poorer hearing was taken into account, which suggests deficits unrelated to audibility. The **group x voicing** cue, emerging when audibility was controlled for, was due to increased effectiveness of voiced cues for YA-NH compared to the other two listener groups.
   PERS: performance was similar for all three groups in noise and in quiet; only effect of group was and interaction with voicing. It occurred because OA used voiced cues more effectively than YA-HI. None of the group effects were affected by audibility.

Next steps

Relating cue perception to cue production on an individual listener-by-listener basis.
Adding children's data to obtain a full picture of cue development and cue use across the lifespan.