

Chapter 5


Age effects on linguistic prosody in coordinates produced to varying interlocutors: Comparison of younger and older speakers

 Clara Huttenlauch^a,  Marie Hansen^a,  Carola de Beer^{a,b},
 Sandra Hanne^a &  Isabell Wartenburger^a

^aUniversity of Potsdam ^bUniversity of Bielefeld

This production study builds on and extends the research on how prosodic cues can be used to resolve syntactic ambiguities. We compared how younger speakers (mean age 25 years, Huttenlauch et al. 2021) and older speakers (mean age 68 years) produced prosodic cues to distinguish between structurally different coordinated three-name sequences without and with internal grouping of the first two names. The prosodic cues of interest were variations in f₀ (F₀ range), duration of segments at the end of the names (final lengthening), and pause insertion. In line with the Proximity/Similarity model by Kentner & Féry (2013), we found that both age groups used all three cues to signal the grouping: Prosodic cues were modified on the group-internal Name1 as well as on Name2 at the right-most element of the group. These prosodic cues were clearly understood by naïve listeners. Successful prosodic disambiguation was not affected by age-related differences in speech production. Furthermore, we analysed the productions with regard to different contexts, such as addressing interlocutors of different ages and mother tongues, and in noisy environments. We found that both age groups of speakers used the same prosodic cues consistently across all contexts, indicating that the use of prosodic cues to clarify syntactic ambiguities is a stable part of the production process, which we interpret as being in line with models of situational independence of disambiguating prosody (e.g., Schafer et al. 2000). Our study provides evidence that the use of these prosodic cues (F₀ range, final lengthening, and pause) is a reliable way to clarify ambiguous structures in speech and independent of the speaker's age.



Clara Huttenlauch, Marie Hansen, Carola de Beer, Sandra Hanne & Isabell Wartenburger. 2023. Age effects on linguistic prosody in coordinates produced to varying interlocutors: Comparison of younger and older speakers. In Fabian Schubö, Sabine Zerbian, Sandra Hanne & Isabell Wartenburger (eds.), *Prosodic boundary phenomena*, 157–192. Berlin: Language Science Press. DOI: 10.5281/zenodo.7777534 

1 Introduction

Linguistic prosody, as in prosodic boundaries, can be used to resolve syntactic ambiguities. Such syntactic ambiguities exist in coordinated sequences of more than two elements (e.g., names) since those elements can be grouped internally at different levels. For instance, the three-name sequence *Moni and Lilli and Manu* can describe three individual persons or a group of three persons (i.e., no internal grouping as in (1)) or a group of two persons in addition to one individual person, with two different possibilities for the grouping (i.e., the group can consist of *Moni and Lilli* or of *Lilli and Manu*. (2) gives an example for the internal grouping of *Moni and Lilli* indicated by parentheses). The latter two different groupings correspond to underlying syntactic structures that differ in their direction of embedding. The difference to the first sequence is the depth of embedding. The absence or type of internal grouping as in (1) versus (2) in an answer to the question *Who will plant a tree?* results in either one, or two, or three planted trees. Prosody, thus, brings the underlying structure to the surface (i.e., disambiguates the otherwise ambiguous surface structure). In this study, we will compare productions of a structure without internal grouping (1) to a structure with internal grouping of the first two elements (2).

- (1) Name1 and Name2 and Name3. – without internal grouping
- (2) (Name1 and Name2) and Name3. – with internal grouping

1.1 Prosodic marking in coordinate sequences

In German, the difference between the two structures (i.e., the resolution of the structural ambiguity) is mainly indicated by one or more of three prosodic cues: F0 change, final lengthening, and pause (Peters et al. 2005, Gollrad et al. 2010, Kentner & Féry 2013, Petrone et al. 2017, for final lengthening see also Schubö & Zerbian 2023 [this volume]). Young speakers have been shown to use these three prosodic cues to clearly mark the internal grouping of coordinated name sequences (Kentner & Féry 2013, Petrone et al. 2017, Huttenlauch et al. 2021). Figure 1 provides visualisations of waveform and spectrogram with F0 contour and segmental annotations of productions without and with internal grouping, respectively, generated using Praat (Boersma & Weenink 2019). The marking of the internal grouping appears as a global and not a local phenomenon, in accordance with the Proximity/Similarity model (Kentner & Féry 2013): Young speakers modified prosodic cues not only at the right edge of the internal group (i.e., on

Name2 in the example in (2)), but already earlier in the utterance (i.e., on Name1, see also left and right panel in Figure 1, Kentner & Féry 2013, Huttenlauch et al. 2021). The principle of Proximity relates to the syntactic constituent structure (Kentner & Féry 2013). The proximity of syntactically grouped elements is expressed by a weakening of the prosodic cues (e.g., less final lengthening, lower F0 peak, smaller F0 range) on the left-most element of two sister elements (e.g., Name1 in (2), Moni in right panel of Figure 1) compared to an ungrouped element in the same position (e.g., Name1 in (1), Moni in left panel of Figure 1). The principle of Anti-Proximity predicts a strengthening of the prosodic cues (e.g., more final lengthening, higher F0 peak, larger F0 range, insertion of a pause) on/after the right-most element of a group than on/after an ungrouped element (e.g., Name2 in (2) versus in (1), Lilli in right versus left panel of Figure 1). The principle of Similarity relates to the depth of syntactic embedding and since it does not apply to our structures we will not discuss it further. In summary, in name sequences with grouping such as (2), the productions of Name1 contain weaker prosodic cues and those of Name2 encompass stronger prosodic cues compared to name sequences without grouping such as (1).

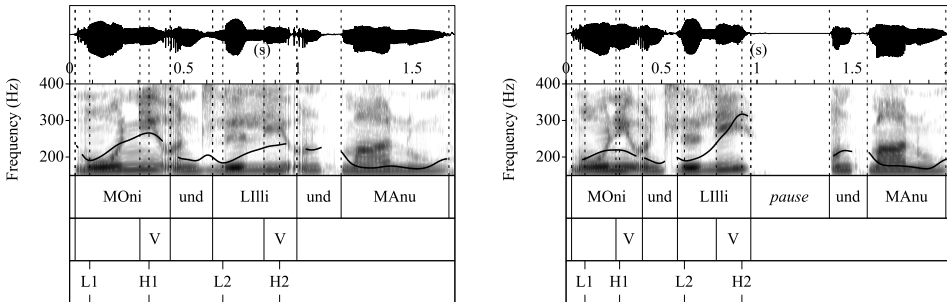


Figure 1: Waveform and spectrogram with F0 contour (black line) of the coordinated name sequence *MOni und Lilli und MAnu* (capital letters correspond to stressed syllable) produced without internal grouping (left) and with internal grouping (right) by a young female speaker. The TextGrid gives an example for the manual annotation of low (L) and high (H) F0 values and the segmentation of the final vowels within Name1 and Name2.

In perception, the early cues on Name1 could reliably be recovered to predict the upcoming structure by more than half of the participants in a two-alternative forced choice decision task with gated stimuli (Hansen et al. 2022). Although all young speakers in Huttenlauch et al. (2021) reliably marked the constituent grouping of coordinated names, they showed inter-speaker variability in how they phonetically realised the prosodic boundary, especially final lengthening

was used in a more flexible way than F0 range and pause. Besides prosodic disambiguation, Huttenlauch et al. (2021) investigated the situational (in)dependence of disambiguating prosody by comparing prosodic cues addressed to interlocutors differing in age and mother tongue as well as in the absence/presence of background noise. Despite the phonetic variability in the realisation of prosodic cues between speakers, the data show a rather consistent pattern of prosodic cues across different communicative situations. The latter finding was interpreted as indexing situational independence: Disambiguating prosody seems to be produced automatically by the speakers in a rather invariant manner.

The present study builds on and extends the results on prosodic boundary production of young speakers (Huttenlauch et al. 2021) with productions of older speakers. Data of both age groups were elicited with the same design and materials, which allows for a direct comparison and detailed investigation of age effects. Age has not only been shown to affect language production in terms of word-finding abilities (for a review see Burke & Shafto 2004) but also in terms of altered acoustic characteristics affecting prosody-related features in the tonal and durational domain. Age, thus, has an effect on the same features that are relevant for the realisation of linguistic prosody.¹ Age, therefore, may interact with the modulation of prosodic cues in conveying the intended meaning. In the remaining part of the introduction, we will address age-related changes in the tonal and durational domain in general (Section 1.2) and their possible impact on the use of linguistic prosody in particular (Section 1.3). Finally, we will present findings on the situational (in)dependence of prosodic cues (Section 1.4).

1.2 Age-related changes in the tonal and durational domain in general

In the following section, we will summarise previous research on general age-related changes in the tonal and durational domains. It is important to note that studies differ in how they group participants into age ranges and in how many years each age group spans. We will use *young* or *younger speakers* to refer to the age range between 18 and 30 years of age and *older speakers* for ages above 60 years.

In the tonal domain, age effects on fundamental frequency (F0) have been studied for several measures including mean and median F0, the span between minimum and maximum (F0 range), and the variability of those measures captured in standard deviations (SD). Here, we focus on the latter two as mean or median F0 are rather uninformative in the context of our study, which focuses on

¹We are aware of the multitude of non-linguistic information transmitted through prosodic cues including but not limited to the emotional state and background of the speaker. In the context of this study, we are only interested in linguistic prosody.

analysing F0 range. So far, results are inconclusive and in part divergent between genders. For F0 range, some studies report no differences between younger and older speakers (Markó & Bóna 2010, Smiljanic & Gilbert 2017), while Dimitrova et al. (2018), Tuomainen & Hazan (2018), and Hazan et al. (2019) observed a larger F0 range for older compared to younger women and Kemper et al. (1998) found a smaller F0 range in older compared to younger speakers irrespective of gender. When it comes to F0 variability, there is evidence for an increase with increasing age (Scukanec et al. 1992, Lortie et al. 2015, Santos et al. 2021). More variability and less stability in older speakers compared to younger speakers was further noticed by several studies looking at more specific measures regarding speech acoustics (including jitter, shimmer, and noise-to-harmonics-ratio; Goy et al. 2013, Lortie et al. 2015, Rojas et al. 2020 among others).

In the durational domain, previous studies observed slower speaking/articulation rates in older compared to younger speakers (Tuomainen & Hazan 2018, Hazan et al. 2019, Tuomainen et al. 2019, 2021 and references in a review by Tucker et al. 2021: 5), relating this finding mainly to longer syllable or word durations (Scukanec et al. 1996, Harnsberger et al. 2008, Barnes 2013, Dimitrova et al. 2018), longer segment durations (Kemper et al. 1995, Harnsberger et al. 2008, Smiljanic & Gilbert 2017), or an increased number of pauses (Kemper et al. 1998, Dimitrova et al. 2018). However, no evidence for pause duration as a driver of age-related differences in speech rate has been reported so far (Barnes 2013, Smiljanic & Gilbert 2017, Dimitrova et al. 2018).

To sum up, previous researchers provided some evidence for tonal and durational differences between younger and older speakers, indicating increased F0 ranges and durations with increasing age. Since these changes affect the same channel used to convey linguistic meaning, we will address possible interferences in the next paragraph.

1.3 Age-related changes in the tonal and durational domain alongside linguistic prosody

We will now turn towards studies that can help to address the question of whether age-related changes in the tonal and durational domain interact with the modulation of disambiguating prosodic cues, as these studies used speech material that explicitly required the use of linguistic prosody. Scukanec et al. (1996) measured the maximal F0 value within the vowel of elicited monosyllabic words in either contrastive or non-contrastive stress position in younger and older female English speakers. Both age groups used F0 in a similar way to mark the focused words (Scukanec et al. 1996: 235). However, independent of the

word position in the sentence, older speakers produced higher F0 values than young speakers in words with contrastive stress and lower maximal F0 values in words in non-contrast positions. The authors concluded that, for the analysed data set, age did not influence the productions of “linguistically salient variations in prosodic output” (Scukanec et al. 1996: 238). The difference in the maximal F0 values between words with and without contrastive stress was even larger in older than in young speakers. The same holds true for the durational domain: Even though older speakers produced longer word durations together with larger standard deviations (i.e., more variability), both age groups used duration to linguistically distinguish stressed from unstressed words.

Further evidence that older speakers use lengthening for prosodic disambiguation despite an overall age-related slower speaking rate comes from Tauber et al. (2010) and Barnes (2013) who reported longer durations for older English speakers in disambiguating contexts. Barnes (2013) elicited structurally ambiguous sentences with either high or low attachment of the prepositional phrase (e.g., *The girl hit the boy with the fan*) in younger and older English speakers. Although the study found longer durations of the direct object and the prepositional phrase regardless of target in the productions of older speakers than in the productions of younger speakers, the overall results revealed that both age groups used the prosodic cues mean F0, pause duration, word duration, and mean intensity similarly to disambiguate ambiguous sentences. However, in another task tapping production of lexical stress to differentiate noun-verb pairs with strong-weak and weak-strong stress patterns, “older adults utilised F0 to a significantly greater extent than young adults” (Barnes 2013: 43). Tauber and colleagues elicited structurally ambiguous sentences (e.g., *The lake froze over a month ago*) to explicitly test for age differences in the realisation of disambiguating prosody in English sentences (Tauber et al. 2010). They found that intonational boundaries (defined as pause duration plus duration of the critical word at the boundary) were longer in older than in younger speakers. Notably, both age groups seem to have had difficulties with the task, as the percentage of sentences which were successfully disambiguated via prosody was 66% for older speakers (above chance, $p < 0.05$) and 59% for the younger age group (not significantly above chance, $p > 0.06$) (Tauber et al. 2010).

In summary, even though age leads to changes in the tonal and temporal domain in general, there is evidence from English speakers that the modulation of prosody to convey linguistic meaning remains unaffected. Older participants even appear to produce prosodic cues in a more extreme way than younger speakers. To the best of our knowledge, there is no study that addressed age differences in the use of prosody to resolve ambiguities in coordinate structures. If the

findings for English ambiguous sentences are transferable to German coordinate structures, we expect that older speakers disambiguate coordinate structures using more extreme prosodic cues than young speakers. This motivates our first research question:

RQ1: Prosodic disambiguation of coordinate name sequences: Do older speakers compared to young speakers show a more extreme use of the three prosodic cues F0 range, final lengthening, and pause on Name1 and Name2 to mark the internal grouping of coordinates in German?

1.4 Situational (in)dependence of prosodic cues

In the remaining part of the introduction, we will address the situational (in)dependence of prosodic cues, a second topic investigated in Huttenlauch et al. (2021). It deals with the effects of different types of interlocutors and the absence/presence of noise on the use of disambiguating prosodic cues. Huttenlauch et al. (2021) compared the use of prosodic cues in five *contexts* involving four female interlocutors: a young adult (YOUNG), a child (CHILD), an elderly adult (ELDERLY), and a young non-native speaker of German (NON-NATIVE) and in noise (the young adult with background white noise, NOISE). The productions directed at the young adult native speaker (i.e., the context YOUNG) were taken as a baseline for comparisons. The findings showed stability in the use of prosodic cues for disambiguating the internal structure of coordinates. That is, individual speakers produced a limited set of cue patterns with only slight shifts in cue distribution across different contexts. This stability in prosodic patterns for disambiguation irrespective of the context was interpreted in favour of models of situational independence of disambiguating prosody (Schafer et al. 2000, Kraljic & Brennan 2005, Speer et al. 2011). These models predict that disambiguating prosody is produced in an automatic way, for the sake of the speakers themselves, and hence depends neither on the presence or absence of an interlocutor, nor on the type of interlocutor or situational setting (e.g., background noise). Despite arguing for situational independence of disambiguating prosody, Huttenlauch et al. (2021) found slight prosodic modifications in the data that can be attributed to context effects. Similarly, as discussed for the prosodic marking of internal grouping of coordinates in the first part of the introduction, the question arises whether age effects in the tonal and durational domain have an impact on the use of F0 range, final lengthening, and pause when speaking in different contexts and whether we find age effects in the situational (in)dependence of prosodic disambiguation. Research on age effects in speech production to different interlocutors is, to our

knowledge, still scarce. In the following, we will briefly summarise existing findings including the context effects found in the productions of young speakers in Huttenlauch et al. (2021).

With regard to addressing a child interlocutor, we will refrain from summarising the immense body of literature treating speech towards preverbal infants since the use of prosody for disambiguation requires that language ability has already been acquired to a certain extent. We are not aware of studies investigating effects of speaker age on prosodic cues uttered towards a child interlocutor. For young speakers, speech towards a child interlocutor has been described as containing an increased F0 range (Biersack et al. 2005, Huttenlauch et al. 2021), lengthened vowels (Biersack et al. 2005), or more pauses (DePaulo & Coleman 1986).

Speech addressing an elderly interlocutor has been explored in data on young and older adult speakers. While younger speakers slowed down their speaking rate by increasing vowel duration and inserting more pauses in speech addressing an elderly interlocutor, older speakers did not do so (Kemper et al. 1995). For older speakers addressing a young interlocutor, however, Kemper and colleagues observed a slower speaking rate than for young speakers. The authors argued that, in comparison to young speakers, older speakers adopt a more simplified speech style including lower speaking rate when addressing a young interlocutor, and thus it is possibly hard for them to slow down even further in order to adapt to an elderly interlocutor (Kemper et al. 1995: 56). Furthermore, young speakers addressing an elderly interlocutor, slowed down their speaking rate with longer pauses, increased final lengthening (Huttenlauch et al. 2021), and increased F0 range or variation in F0 (Thimm et al. 1998, Huttenlauch et al. 2021).

We are not aware of studies investigating effects of speaker age on prosodic cues when addressing a non-native interlocutor. Some studies involving young speakers found no clear differences (DePaulo & Coleman 1986, Uther et al. 2007, Knoll & Scharrer 2007, Knoll et al. 2011, Huttenlauch et al. 2021), while others observed a lowered speech rate due to lengthened pauses (Biersack et al. 2005), a higher mean F0 (Knoll et al. 2015), increased word durations and intensity (Rodriguez-Cuadrado et al. 2018), or an increased F0 range along with segmental modifications described as a more emphatic style (Smith 2007; see Piazza et al. 2021 for a review on foreigner-directed speech).

Finally, speech in noisy environments compared to silent environments is affected by modulations in several ways. The reported changes are referred to as “Lombard speech” (Lombard 1911 as cited in Zollinger & Brumm 2011) and include decreased speaking rate (due to increased segment or word durations), increased

F0 ranges, increased signal amplitude, and spectral changes such as smaller spectral slope (e.g., Junqua 1996, Summers et al. 1988, Jessen et al. 2003, Zollinger & Brumm 2011, Smiljanic & Gilbert 2017, Tuomainen et al. 2019, 2021). The findings for young speakers in a noisy environment in Huttenlauch et al. (2021) were interpreted as being partly in line with Lombard speech, as they revealed increased final lengthening and decreased pause duration but no changes in F0 range. With respect to age effects in speech adaptation to noise, no age differences were found by Dromey & Scott (2016) and Smiljanic & Gilbert (2017), with the latter reporting an age-independent decrease in speaking rate when noise was present, while Tuomainen et al. (2019) reported a decreased speaking rate only for the older age group.

To summarise, the modifications of prosodic cues in coordinates induced by varying contexts observed by Huttenlauch et al. (2021) were rather small but in line with previous findings. The effect of age on the realisation of prosodic cues in more communicative settings with varying interlocutors is still only scarcely explored. For the reported age-related changes in addressing different interlocutors, the question arises whether they replicate to coordinate structures in German. Given the limited evidence, we keep our second research question rather open:

RQ2: Situational (in)dependence: Do young and older speakers differ in adapting their use of prosodic cues when addressing varying interlocutors?

In the current study, we extend the age range of usually studied participants (in Huttenlauch et al. 2021 19–34 years) to older people aged between 60 and 80 years of age (i.e., comparable to the older age groups in the previously presented literature) and compare the productions of linguistic prosody in young and older adult speakers. Specifically, we explore whether age interacts with the modulation of prosodic cues, especially F0 range, final lengthening, and pause, and whether any such interaction may impact the disambiguation of structurally ambiguous coordinated name sequences and the use of prosodic cues when addressing different interlocutors (i.e., regarding situational (in)dependence of disambiguating prosody).

2 Methods and material

Methods, materials, and data of the younger speakers are taken from Huttenlauch et al. (2021) and extended by the data of older speakers.

2.1 Participants

Fifteen young monolingual German native speakers (13 female, 1 male, 1 other; age range: 19–34, mean 25.47 years, SD: 4.6; see Huttenlauch et al. 2021) and 13 older monolingual German native speakers (9 female, 3 male, 1 no information; age range: 61–80 years, mean: 67.77 years, SD: 6.8) were included in the study. Additional five speakers took part in the study, but were discarded due to low task compliance ($n = 1$), scores below 25 in the Montreal Cognitive Assessment (Nasreddine et al. 2005) ($n = 3$), or missing data ($n = 1$). All participants (henceforth *speakers*) were recruited in Potsdam, Germany, and were reimbursed or received course credits (the latter only applies to the young speakers). They were naïve to the purpose of the study and gave written consent to participate. The Ethics Committee of the University of Potsdam approved the procedure of this study (approval number 72/2016). Hearing ability was assessed by a hearing screening using an audiometer (Hortmann DA 324 series) and calculated following the grades of hearing impairment by the WHO as reported in Olusanya et al. (2019). Normal hearing was defined as an average pure-tone audiometry of 25 dB HL or better of 500, 1000, 2000, and 4000 Hz in the better ear. Following this definition, all 15 young speakers and 10 of the older speakers had normal hearing, the remaining speakers showed a slight ($n = 2$) or moderate impairment ($n = 1$).

2.2 Stimuli

2.2.1 Items

As stimuli, we used the same six coordinated name sequences as in Holzgrefe-Lang et al. (2016), Huttenlauch et al. (2021), and Wellmann et al. 2023 [this volume]: Each sequence consisted of three German names coordinated by *und* (English ‘and’) that appeared in each of two conditions: without internal grouping (3) or with internal grouping of the first two names (4). The grouping of the first two names was visually indicated to the participants by bracketing Name1 and Name2 with parentheses as in (4). The conditions will henceforth be referred to as *brack* for the condition with internal grouping and *nobrack* for the condition without internal grouping. A total of 12 items was used. Young speakers produced each item once per context (see Section 2.2.2), older speakers twice to enlarge the data set and to increase statistical power.

- | | | |
|-----|------------------------------|----------------------------|
| (3) | Name1 and Name2 and Name3. | Moni und Lilli und Manu. |
| (4) | (Name1 and Name2) and Name3. | (Moni und Lilli) und Manu. |

The set of coordinates comprised nine different German names in total, all of which were controlled for number of syllables (disyllabic), stress pattern (penultimate), and sonority of the segments (only sonorant material, to facilitate pitch tracking). Six of the names featured the high frontal vowel /i/ in word-final position (Moni, Lilli, Leni, Nelli, Mimmi, and Manni) in order to decrease glottalisation and occurred as Name1 or as Name2. Name3 contained either /u/ or /a/ in word-final position (Manu, Nina, and Lola). Regarding possible collocations of the selected names for each coordinate, there was no particular co-occurrence of two adjacent names (as in, e.g., “Bonnie and Clyde”) in the dlexDB corpora (Heister et al. 2011) or in printed sources between 1500 and 2021, as ascertained by the Google Ngram Viewer (Lin et al. 2012).

2.2.2 Contexts

Five different communicative contexts (YOUNG, CHILD, ELDERLY, NON-NATIVE, NOISE) were created that differed in the interlocutor and/or the absence/presence of background white noise (see Table 1). Speakers saw their interlocutors on a screen in two short videos each (one with a personal introduction of the interlocutor and one with instructions for the task) to get an audio-visual impression. The young and non-native interlocutors were similar in age to the group of young speakers, the elderly interlocutor was two years older than the oldest speaker in the group of older speakers. A more detailed description of the videos and interlocutors can be found in Huttenlauch et al. (2021).

2.3 Procedure

Productions were elicited by means of a referential communication task. Contexts were presented blockwise, always starting with the YOUNG context, which served as a baseline in the analysis. The order of the other four contexts was randomised. Each block started with the two video clips of the corresponding interlocutor. Then, for each trial, speakers first saw a fixation cross on the screen accompanied with the auditory presentation of the trigger question *Wer kommt?* (‘Who is coming?’) via headphones produced by the interlocutor of the current block as a reminder to whom they were talking. After 1000 ms, the fixation cross was replaced by the visual presentation of the name sequence (i.e., the item) in one of the two conditions (see Figure 2). The task was to produce the item in a way that would allow the interlocutor “to understand as rapidly and accurately as possible who is coming together”. Recordings took place in a sound-attenuated booth at the University of Potsdam via an Alesis iO2 audio interface using an AKG HSC271 headset with over-ear headphones and a condenser microphone.

Table 1: Fictional names, ages, origins, and further information of the interlocutors present in the five contexts.

	YOUNG (baseline)	CHILD	ELDERLY	NON-NATIVE	NOISE
Name:	Hannah	Carlotta	Maria Korbmacher	Zsófi	Hannah + white noise
Age (in years):	24	6	82	26	See YOUNG
Origin:	Eberswalde	Potsdam	NA	NA	
Residence:	Potsdam	Potsdam	Potsdam	Potsdam	
Occupation:	Biology student	School child	Retired school teacher	Exchange student	
Further facts:	Moved to Potsdam for her studies, lives in a shared flat, likes the parks in Potsdam	Likes horse riding, her parents pick her up from school, is good at swimming	Lives for two years in an old-age home with her husband, tends to forget things from time to time	Started to learn German one year ago, lives in a shared flat, enjoys doing sports	

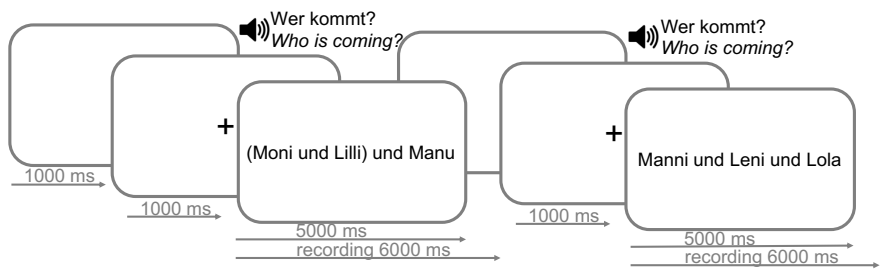


Figure 2: Experimental setting and timing of two trials.

The wide screen in the recordings booth had a resolution of 1920×1200, stimuli were written in Arial, font size 50. The experiment was run from a Dell laptop using Presentation software (Neurobehavioural Systems 2018). Each item was presented in each context once (for young speakers) or twice (for older speakers). Thus, the data set contained 900 individual productions of young speakers (6 name sequences × 2 conditions × 5 contexts × 15 young speakers) and 1560 individual productions of older speakers (6 name sequences × 2 conditions × 5 contexts × 2 repetitions × 13 speakers).

2.4 Perception check

After data collection of the production study, all recordings were auditorily presented to naïve listeners who were asked to indicate for each production the perceived condition. To this end they were given two pictograms with three persons each, one pictogram per condition (Figure 3, picture A without and picture B with internal grouping).

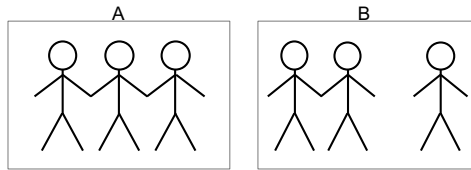


Figure 3: Pictograms used in the perception check depicting the condition without grouping (left panel) and with grouping (right panel).

The aim of the perception check was to assess whether naïve listeners perceive the grouping of the coordinates in the way it was *intended*. By *intended* we refer to the indication of condition which was given to speakers by parentheses around the grouped names in the production study. Obviously, the intention of speakers at the time of the production remains unknown to us.

The data of the young and older age groups were rated separately. The recordings were distributed across different lists with 147 to 267 items. Each listener judged one list and each list was judged by seven or eight listeners.

The perception check of the productions of the young speakers was conducted in presence of several listeners in the same room with a paper-and-pen version. Data of 31 listeners (22 female, 9 male; age range: 18–41, mean: 24.1 years, SD: 5.8) were analysed. Another 11 listeners took part in the study, but had to be excluded due to technical problems ($n = 9$), German as a non-native language ($n = 1$) or a hit-rate 2 SD below the mean hit-rate of all listeners ($n = 1$, see Huttenlauch et al. 2021 for more details).

For the productions of the older speakers, the perception check was transferred onto OpenSesame (Mathôt et al. 2012) and was run as a web-based study on JATOS (Lange et al. 2015) in individual sessions. Data of 49 listeners (29 female, 9 male, 11 other/no information; age range: 18–63, mean: 24.63 years, SD: 6.3) were analysed. Another five listeners took part in the study, but had to be excluded due to technical problems.

In the analysis of the perception check, the exclusion threshold for individual productions was set to a hit-ratio 2 SD below the mean ratio, as suggested by standard assumptions on the exclusion of data points (e.g., Howell et al. 1998). Hit-ratio was calculated separately for each production as the number of congruent rates (i.e., correct identification of the intended grouping/condition, referred to as *hit-rate*) divided by the number of total rates. Applying this criterion, 36 productions (4%, 11 nobrack, 25 brack) in the group of the young speakers and 66 productions (4%, 39 nobrack, 27 brack) in the group of the older speakers fell below the threshold and were excluded from further analyses. For a more detailed description of procedure and analysis of the perception check see Huttenlauch et al. (2021).

2.5 Segmentation and measurements

In addition to the productions excluded based on the perception check, three productions were excluded from analysis in the data set of the older speakers: due to hesitations that made the analysis of condition impossible ($n = 2$) and due to recording problems ($n = 1$). The final data set comprised 2355 productions (young: 864, older: 1491). Table 2 provides an overview of how the productions distribute across age groups, conditions, and contexts.

Table 2: Distribution of productions entering statistical analyses across age groups, conditions, and contexts in the final data set.

age group	condition	YOUNG	CHILD	ELDERLY	NON-NATIVE	NOISE
younger	nobrack	87	85	90	90	87
	brack	83	88	89	86	79
older	nobrack	141	148	148	151	153
	brack	151	153	153	148	145

For the extraction of the three prosodic cues under investigation, segment boundaries and pauses were manually annotated in Praat (Boersma & Weenink 2019, version 6.0.32) by following the criteria in Turk et al. (2006). Silent intervals

of at least 20 ms duration were considered as pauses (following the procedure in Petrone et al. 2017). F0-minima (L) and F0-maxima (H) on both Name1 and Name2, were manually annotated (example TextGrids are given in Figure 1). The points were set into parts of the signal, where F0 can be reliably measured (i.e., avoiding the edges of segments, glottalised parts in the signal, and parts with other non-modal voice quality). The F0 contour mostly displayed a rising movement on Name1 and Name2, respectively (i.e., L preceded H). Only in a few cases, speakers produced a falling F0 movement on Name1 (young speakers: 88 falls versus 776 rises, older speakers: 108 falls versus 1368 rises) or Name2 (older speakers: 13 falls versus 1458 rises). For some productions in the data of the elderly speakers it was impossible to find reliable locations to annotate either L and/or H points and it was, thus, impossible to measure the F0 range. In those cases, the corresponding item was excluded from the analysis of F0 range for Name1 and/or Name2. This applies to 15 items (1.0% of the productions of older speakers) in the condition without internal grouping and to 20 items (1.3% of the productions of older speakers) with internal grouping. All in all, we aimed for an approach of measuring F0 range that was applicable to the majority of the recordings. For further segmentation criteria see Huttenlauch et al. (2021). For Name1 and Name2 separately, we calculated the three variables F0 range, final lengthening, and pause. The variable F0 range reflects the range between the F0-minimum and the F0-maximum on NameX in semitones (st; calculated as $12 \times \log_2(F0_H/F0_L)$). The variable final lengthening reflects the duration of the final vowel of NameX divided by the duration of NameX (in %, the final vowel is annotated as V on the second tier of the TextGrid in Figure 1). The pause variable reflects the duration of a possible pause after NameX divided by the duration of the whole utterance (in %). We chose relative instead of absolute measures as they are independent of individual speech rates and mean fundamental frequency. However, to descriptively assess potential age-related effects, absolute durational measurements were taken into consideration.

2.6 Statistical analysis

The workflow of the statistical analyses was similar to that in Huttenlauch et al. (2021), additionally comprising a group comparison between young and older speakers. For each dependent variable (F0 range, final lengthening, pause) on Name1 and Name2, we ran separate linear mixed-effects regression models in R (R Core Team 2018). Each model estimated the difference in the dependent variables between the two age groups (young and older speakers), between the four context comparisons, and between the two conditions (brack and nobrack), if applicable. Interactions between context and age group were added to further

explore the dependencies of the differences, as well as interactions of context and age group with condition. A maximal model including all main effects and their interactions, as previously described, as well as including a random effects structure with all possible variance components and correlation parameters associated with the four within-subject contrasts (CHILD vs. YOUNG, ELDERLY vs. YOUNG, NON-NATIVE vs. YOUNG, NOISE vs. YOUNG) was always fit first.² In order to avoid overfitting of the random effects structure, we followed the approach outlined in Bates et al. (2015) and conducted an iterative reduction of model complexity. A more detailed explanation of the model reduction, along with all reduced models and the complete model outputs of the fixed effects, can be found on an Open Science Framework project page (<https://osf.io/fc8nz>) together with the data and code. In the results section, we will only report the statistically significant effects which comprise main effects of condition and/or main effects and interactions of age group.

3 Results

In the following, we will first present descriptive results from absolute and relative measurements with a focus on age, including a statistical comparison of the age groups. Hereafter, we will turn towards the results of linear mixed models fit to compare the age groups regarding their use of prosodic cues for disambiguation (RQ1) and regarding their adaptation to different interlocutors (RQ2).

3.1 Descriptive statistics and statistical age group comparison of absolute durational measurements

In the main section of our analysis, we analysed the use of prosodic cues by measuring the relative duration of speech segments and pauses. This method allowed us to understand how prosodic cues were used, regardless of individual

²Prosodic_cue ~ 1 + condition*context*age_group +
(1 + condition +
child_vs_young + elderly_vs_young + nonnat_vs_young + noise_vs_young +
age_group +
condition:age_group +
condition:child_vs_young + condition:elderly_vs_young +
condition:nonnative_vs_young + condition:noise_vs_young +
child_vs_young:age_group + elderly_vs_young:age_group +
nonnative_vs_young:age_group + noise_vs_young:age_group +
condition:child_vs_young:age_group + condition:elderly_vs_young:age_group +
condition:nonnative_vs_young:age_group +
condition:noise_vs_young:age_group | speaker)

differences in speaking rate or the absolute duration of sounds. Before presenting the relative measurements, we will present some absolute durational measurements to compare the differences between younger and older speakers (cf. Table 3). However, we will not include measurements of average F0 by age group because the speaker groups had mixed genders, which could affect our estimation of differences in F0 between the groups.

Table 3: Descriptive statistics of absolute durational measurements by age group and statistical group comparison.

Measurement (ms)	Younger		Older		Comparison
	mean	SD	mean	SD	<i>p</i>
utterance duration	1964.63	292.16	2181.25	444.80	< 0.0001
final vowel duration (Name1)	129.61	40.09	144.68	46.20	< 0.0001
pause duration (after Name2)	172.93	195.24	262.83	330.05	< 0.0001
final vowel duration (Name2)	181.53	59.51	198.24	65.57	< 0.0001

In our data set we observe longer absolute durations for older as compared to younger speakers for the whole utterance (mean difference of 217 ms), the final vowels of Name1 and Name2 (mean difference of 15 ms and 17 ms, respectively), and the pause after Name2 (mean difference of 89.9 ms). All age group comparisons were statistically significant in linear models with age group as a single sum-contrasted predictor (0.5 for young and -0.5 for older speakers). Moreover, we observe a higher degree of variation (larger SDs) for older speakers than for young speakers across all durational measurements.

3.2 Descriptive statistics of relative measurements

Relative measurements of F0 range, final lengthening, and pause were used to explore the use of prosodic cues for the disambiguation of coordinates with and without internal grouping. Figure 4 shows a visual description of mean location and spread of F0 range as well as final lengthening on Name1 by age group, context, and condition. For both cues and for each context, the mean values in the brack condition are lower for younger than for older speakers, while in the no-brack condition in all contexts except YOUNG, the mean values are larger for younger compared to older speakers. Considering these raw data visually, the difference between conditions is larger in the productions of young speakers than in that of older speakers. We did not run statistical analyses and do not

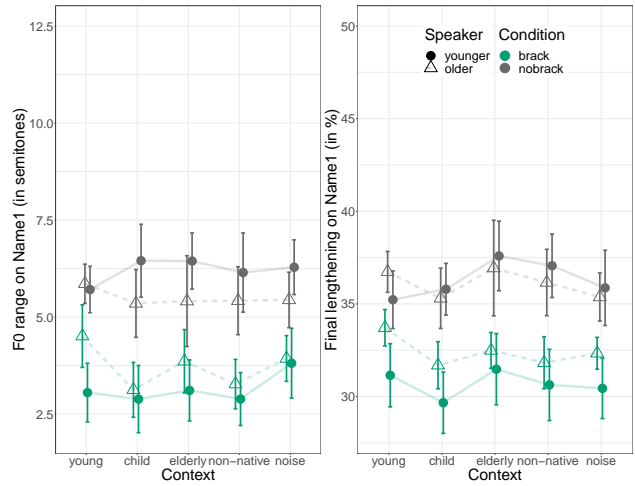


Figure 4: Distribution of raw values of F0 range (left panel) and final lengthening (right panel) on Name1 (y-axis) divided by context (x-axis), condition (colour: grey for nobrack, green for brack), and age group (shape: circles for young speakers, triangles for older speakers). Whiskers show 95% confidence intervals.

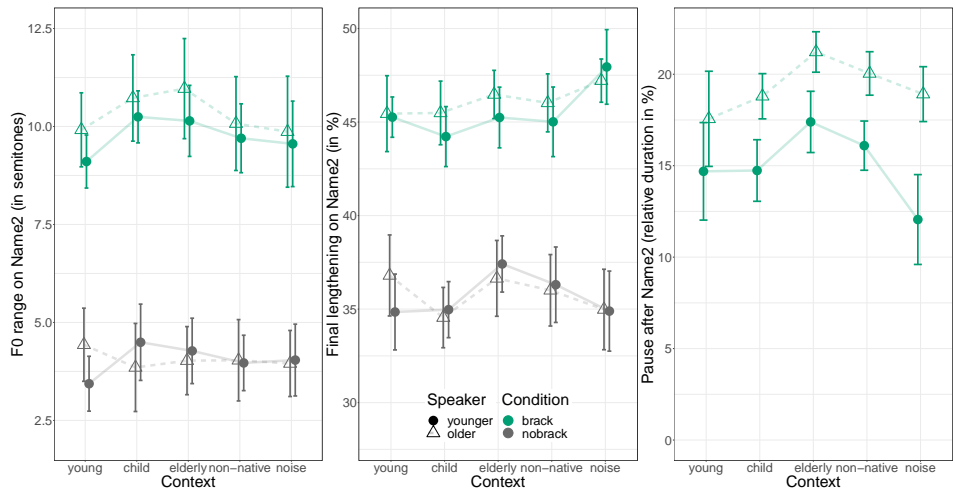


Figure 5: Distribution of raw values of F0 range (left panel), final lengthening (mid panel), and pause (right panel) on Name2 (y-axis) divided by context (x-axis), condition (colour: grey for nobrack, green for brack), and age group (shape: circles for young speakers, triangles for older speakers). Whiskers show 95% confidence intervals.

report descriptive statistics on pause duration after Name1 since mostly zero values were produced by the participants. That is, a pause after Name1 was only produced in 206 out of 2355 trials in total, 175 times in the nobrack condition and 31 times in the brack condition. Figure 5 shows a visual description of mean location and spread of F0 range, final lengthening, and pause on/after Name2 by age group, context, and condition. There is no apparent visual pattern that would apply to both speaker groups and all three cues. For F0 range and pause in the brack condition, young speakers produced smaller mean values than older speakers. For final lengthening in general and F0 range of the nobrack condition, the values are more mixed between age groups. With regard to the direction of the difference in the degree of F0 range and final lengthening between the brack and nobrack condition, both prosodic cues show smaller values in brack than in nobrack on Name1 and the opposite pattern, larger values in brack than in nobrack, on Name2.

To summarise, a visual inspection of the raw data reveals differences between the two age groups in the amount to which the different prosodic cues were produced in the respective contexts and conditions. Nevertheless, the general patterns for each cue are quite similar across contexts for both, young and older speakers. That is, for instance for F0 range in the brack condition in Figure 5 (left panel, green data points), the connecting lines between contexts have slopes in the same directions between speaker groups and in any case do not cross. We are aware that the descriptive analysis of the data does not allow for any generalisations. In the following sections, we will present the results of the statistical models we ran on each cue and Name individually.

3.3 Statistical analyses on Name1

3.3.1 F0 range on Name1

Results for F0 range on Name1 are reported from a reduced model³ (all final models and code can be found on <https://osf.io/fc8nz>). Several effects were statistically significant (see Table 4 and <https://osf.io/fc8nz>).

³F0_name1 ~ 1 + condition*context*age_group +
 (1 + child_vs_young + elderly_vs_young + noise_vs_young +
 age_group +
 condition:age_group +
 nonnative_vs_young:age_group +
 condition:child_vs_young:age_group +
 condition:nonnative_vs_young:age_group | speaker)

Table 4: Selected model estimates and 95% confidence intervals of the fixed effects for F0 range on Name1 including main effect of condition and main effect and interactions of age group. * $p < 0.05$; ** $p < 0.01$.

Predictor	Estimate	95% CI
Intercept	4.666**	(4.060, 5.273)
condition	-1.236**	(-1.559, -0.913)
age group	0.002	(-1.211, 1.216)
condition:age group	-0.593	(-1.239, 0.053)
CHILD vs. YOUNG:age group	1.225**	(0.563, 1.886)
ELDERLY vs. YOUNG:age group	0.757	(-0.259, 1.773)
NON-NATIVE vs. YOUNG:age group	0.928*	(0.217, 1.639)
NOISE vs. YOUNG:age group	1.193*	(0.230, 2.155)
condition:CHILD vs. YOUNG:age group	0.051	(-0.515, 0.616)
condition:ELDERLY vs. YOUNG:age group	-0.013	(-0.450, 0.423)
condition:NON-NATIVE vs. YOUNG:age group	0.130	(-0.404, 0.664)
condition:NOISE vs. YOUNG:age group	0.271	(-0.170, 0.712)

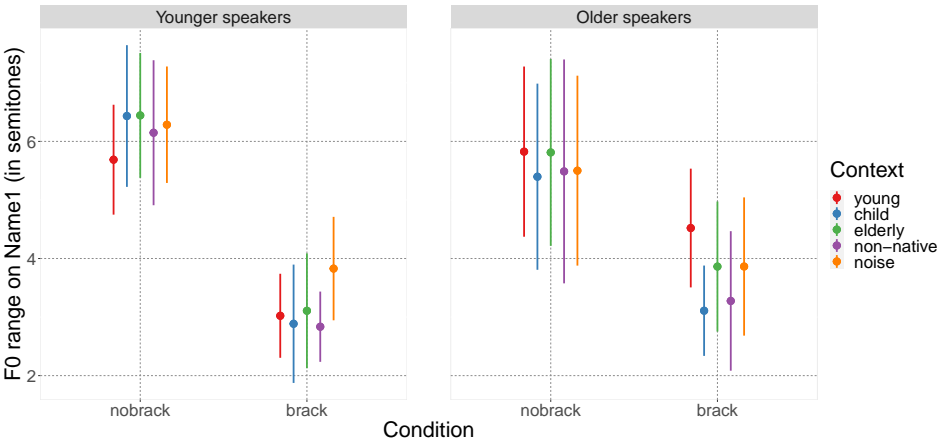


Figure 6: Model predictions for F0 range on Name1 (y-axis) divided by age group (younger speakers left panel, older speakers right panel), condition (x-axis), and context (colour). Whiskers show 95% confidence intervals.

The statistically significant main effect of condition ($\beta = -1.236, p < 0.0001$) confirms that F0 range was used for the disambiguation of brack and nobrack on Name1 by speakers of both age groups: The F0 range in the brack condition was decreased by about 2.5 semitones compared to the nobrack condition. With respect to age-related differences in situational (in)dependence, the statistically significant two-way interactions of the context comparisons CHILD vs. YOUNG ($\beta = 1.225, p = 0.0003$), NON-NATIVE vs. YOUNG ($\beta = 0.928, p = 0.011$), and NOISE vs. YOUNG ($\beta = 1.193, p = 0.016$) with age group, respectively, indicate general age-related differences when addressing the child and non-native as compared to the young interlocutor, as well as age-related differences in noisy vs. non-noisy settings with a young interlocutor. In all of the three context comparisons, young speakers increased their F0 range compared to context YOUNG, while older speakers decreased their F0 range. Model predictions for F0 range on Name1 by condition, context, and age group are displayed in Figure 6.

3.3.2 Final lengthening on Name1

Results for final lengthening on Name1 are reported from a reduced model.⁴ Several effects were statistically significant (see Table 5 and link in Section 2.6). The statistically significant main effect of condition ($\beta = -2.366, p < 0.0001$) confirms that final lengthening was used for the disambiguation of brack and nobrack on Name1 by speakers of both age groups: Final lengthening was decreased in the brack condition (where the final vowel span about 31% of the total name duration) as compared to the nobrack condition (where the final vowel span about 36% of the total name duration). With respect to age-related differences in situational (in)dependence, the statistically significant two-way interaction of the context comparison CHILD vs. YOUNG with age group ($\beta = 1.449, p = 0.002$) indicates that young speakers, in contrast to older speakers, increased final lengthening when addressing the child compared to the young interlocutor. A similar pattern is predicted by the model for the context comparison NON-NATIVE vs. YOUNG, for which the interaction with age group was statistically significant ($\beta = 1.877, p = 0.028$): While final lengthening is increased by young speakers when addressing the non-native as compared to the young interlocutor, final lengthening is decreased by older speakers. Model predictions for final lengthening on Name1 by condition, context, and age group are displayed in Figure 7.

⁴The model can be found at <https://osf.io/fc8nz>.

Table 5: Selected model estimates and 95% confidence intervals of the fixed effects for final lengthening on Name1 including main effect of condition and main effect and interactions of age group. * $p < 0.05$; ** $p < 0.01$.

Predictor	Estimate	95% CI
Intercept	33.848**	(32.716, 34.980)
condition	-2.366**	(-2.949, -1.784)
age group	-0.794	(-3.058, 1.469)
condition:age group	-1.001	(-2.166, 0.164)
CHILD vs. YOUNG:age group	1.449*	(0.063, 2.834)
ELDERLY vs. YOUNG:age group	1.962	(-0.255, 4.179)
NON-NATIVE vs. YOUNG:age group	1.877*	(0.203, 3.551)
NOISE vs. YOUNG:age group	1.371	(-0.289, 3.032)
condition:CHILD vs. YOUNG:age group	-0.841	(-2.226, 0.545)
condition:ELDERLY vs. YOUNG:age group	-0.361	(-1.740, 1.017)
condition:NON-NATIVE vs. YOUNG:age group	-0.612	(-1.995, 0.771)
condition:NOISE vs. YOUNG:age group	-0.726	(-2.124, 0.672)

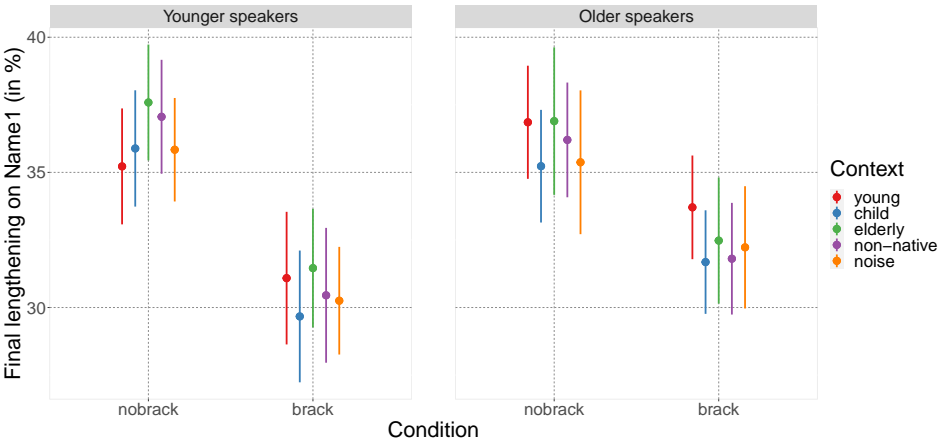


Figure 7: Model predictions for final lengthening on Name1 (y-axis) divided by age group (younger speakers left panel, older speakers right panel), condition (x-axis), and context (colour). Whiskers show 95% confidence intervals.

3.4 Statistical analyses on Name2

3.4.1 F0 range on Name2

Results for F0 range on Name2 are reported from a reduced model.⁵ Several effects were statistically significant (see Table 6 and link in Section 2.6). The statistically significant main effect of condition ($\beta = 3.04$, $p < 0.0001$) confirms that F0 range was used for the disambiguation of brack and nobrack on Name2 across both age groups: The F0 range in the brack condition was increased by about six semitones compared to the nobrack condition. With respect to age-related differences in situational (in)dependence, the significant two-way interaction of the context comparison CHILD vs. YOUNG with age group ($\beta = 0.873$, $p = 0.011$) indicates general age-related differences in approaching the child interlocutor compared to the young interlocutor: The F0 range was larger for young speakers than that of older speakers when addressing the child in comparison to the young interlocutor. These age-related patterns diverge even more when context-related prosodic disambiguation is considered and condition is taken into account. The significant three-way interaction of condition, context comparison CHILD vs. context YOUNG, and age group ($\beta = -0.799$, $p = 0.018$) indicates that young speakers increased the F0 range in both conditions, brack and nobrack, when addressing the child as compared to the young interlocutor, while older speakers did so only in the brack condition. In the nobrack condition, however, older speakers decreased the F0 range, resulting in an enhanced difference between the conditions when addressing the child as compared to the young interlocutor. Model predictions for F0 range on Name2 by condition, context, and age group are displayed in Figure 8.

3.4.2 Final lengthening on Name2

Results for final lengthening on Name2 are reported from a reduced model.⁶ Several effects were statistically significant (see Table 7 and link in Section 2.6). The statistically significant main effect of condition ($\beta = 5.071$, $p < 0.0001$) confirms that final lengthening was used for the disambiguation of brack and nobrack on Name2 by speakers of both age groups: Final lengthening was increased in the brack condition (the final vowel of Name2 span about 45% of the total duration of Name2) compared to the nobrack condition (the final vowel span about 35% of the total name duration). Regarding age-related differences in prosodic disambiguation and situational (in)dependence, the three-way interaction between

⁵The model can be found on <https://osf.io/fc8nz>.

⁶The model can be found on <https://osf.io/fc8nz>.

Table 6: Selected model estimates and 95% confidence intervals of the fixed effects for F0 range on Name2 including main effect of condition and main effect and interactions of age group. * $p < 0.05$; ** $p < 0.01$.

Predictor	Estimate	95% CI
Intercept	7.097**	(6.370, 7.824)
condition	3.040**	(2.613, 3.468)
condition:age group	−0.345	(−1.200, 0.510)
CHILD vs. YOUNG:age group	0.873*	(0.121, 1.626)
ELDERLY vs. YOUNG:age group	0.573	(−0.413, 1.559)
NON-NATIVE vs. YOUNG:age group	0.636	(−0.384, 1.655)
NOISE vs. YOUNG:age group	0.642	(−0.351, 1.635)
condition:CHILD vs. YOUNG:age group	−0.779*	(−1.419, −0.139)
condition:ELDERLY vs. YOUNG:age group	−0.684	(−1.524, 0.156)
condition:NON-NATIVE vs. YOUNG:age group	−0.306	(−0.968, 0.356)
condition:NOISE vs. YOUNG:age group	−0.442	(−1.193, 0.309)

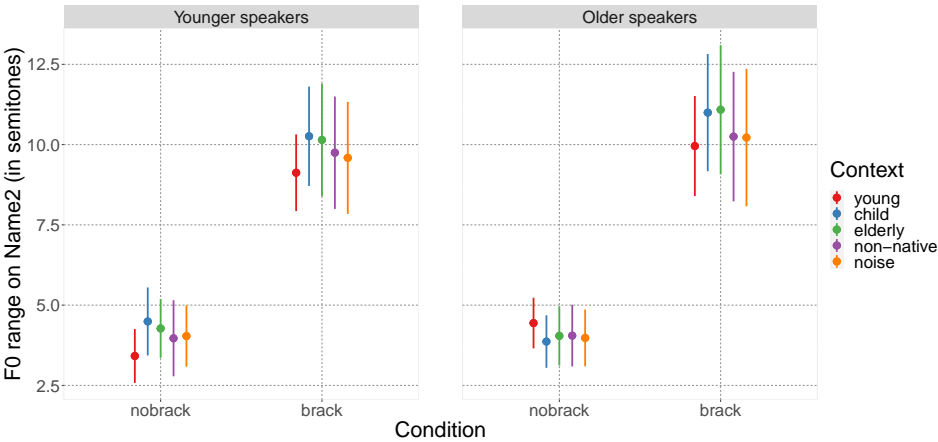


Figure 8: Model predictions for F0 range on Name2 (y-axis) divided by age group (younger speakers left panel, older speakers right panel), condition (x-axis), and context (colour). Whiskers show 95% confidence intervals.

Table 7: Selected model estimates and 95% confidence intervals of the fixed effects for final lengthening on Name2 including main effect of condition and main effect and interactions of age group. * $p < 0.05$; ** $p < 0.01$.

Predictor	Estimate	95% CI
Intercept	40.813*	(39.477, 42.149)
condition	5.071*	(4.284, 5.858)
condition:age group	-0.248	(-1.822, 1.325)
CHILD vs. YOUNG:age group	0.766	(-0.584, 2.116)
ELDERLY vs. YOUNG:age group	0.943	(-0.399, 2.286)
NON-NATIVE vs. YOUNG:age group	0.880	(-0.467, 2.227)
NOISE vs. YOUNG:age group	1.374	(-0.368, 3.115)
condition:CHILD vs. YOUNG:age group	-1.811**	(-3.161, -0.462)
condition:ELDERLY vs. YOUNG:age group	-1.939**	(-3.281, -0.596)
condition:NON-NATIVE vs. YOUNG:age group	-1.455	(-3.035, 0.125)
condition:NOISE vs. YOUNG:age group	-0.536	(-1.898, 0.827)

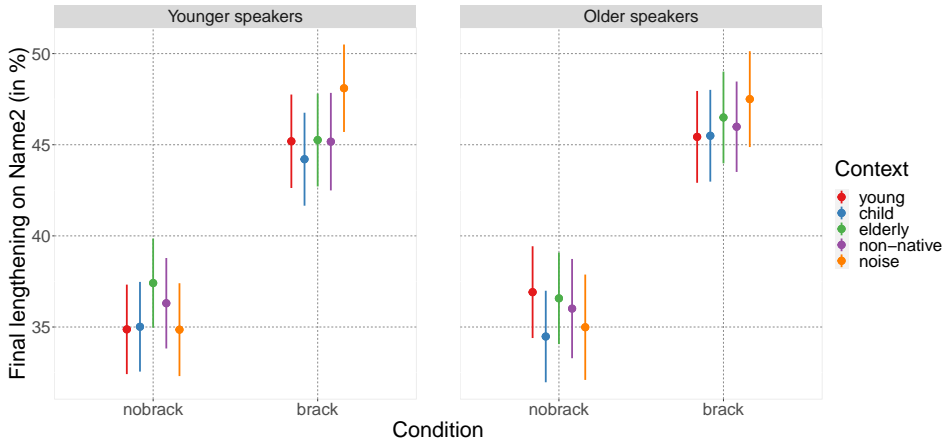


Figure 9: Model predictions for final lengthening on Name2 (y-axis) divided by age group (younger speakers left panel, older speakers right panel), condition (x-axis), and context (colour). Whiskers show 95% confidence intervals.

condition, the context comparison CHILD vs. YOUNG, and age group ($\beta = -1.811$, $p = 0.009$) indicates that young speakers decreased final lengthening in the brack condition when addressing the child as compared to addressing the young interlocutor, thus decreasing the difference between the conditions. On the contrary, older speakers decreased final lengthening for the same context comparison in the nobrack condition, thus increasing the difference between the conditions. An additional three-way interaction between condition, ELDERLY vs. YOUNG and age group ($\beta = -1.939$, $p = 0.005$) indicates that young speakers increased final lengthening in the nobrack condition when addressing the elderly as compared to the young interlocutor. That is, they reduced the difference between the conditions in context ELDERLY, compared to context YOUNG. Older speakers showed a different behaviour: They increased final lengthening when addressing the elderly as compared to the young interlocutor in the brack condition, thus enhancing the difference between the conditions in context ELDERLY. Model predictions for final lengthening on Name2 by condition, context, and age group are displayed in Figure 9.

3.4.3 Pause after Name2

Since the random effects structure of the model analysing pause after Name2 could not be reduced without a significant drop in model fit, results are reported from the maximal model. None of the effects were statistically significant (see Table 8 and link in Section 2.6).

Table 8: Selected model estimates and 95% confidence intervals of the fixed effects for pause after Name2 including main effects and interactions of age group. * $p < 0.05$; ** $p < 0.01$.

Predictor	Estimate	95% CI
Intercept	19.398**	(14.827, 23.968)
age group	-0.019	(-9.160, 9.122)
CHILD vs. YOUNG:age group	-1.560	(-10.505, 7.386)
ELDERLY vs. YOUNG:age group	-1.202	(-10.113, 7.709)
NON-NATIVE vs. YOUNG:age group	-1.048	(-10.001, 7.906)
NOISE vs. YOUNG:age group	16.472	(-21.986, 54.930)

4 Discussion

In the current study, we compared the use of prosodic cues produced to disambiguate the internal grouping of coordinated three-name sequences (coordinates) in two conditions, that is, without and with internal grouping of the first two names (nobrack and brack, respectively) between two age groups: young (19–34 years) and older (61–80 years) speakers of German. We focused our analysis on the three prosodic cues F0 range, final lengthening, and pause on/after Name1 and Name2. As age affects the stability and variability of tonal and durational features in general, we tested for potential age effects on the modulation of the three prosodic cues for structural disambiguation. Furthermore, we explored whether the situational (in)dependence of disambiguating prosody differs between younger and older speakers, considering their prosodic adaptation to varying contexts. To this end, in both age groups, we elicited coordinates by means of a referential communication task with five contexts: addressing a young adult, a child, an elderly adult, a young non-native adult, and the young adult with background noise.

Looking at the data, we note two things: First of all, descriptively, younger and older speakers produced the three prosodic cues overall quite similarly for prosodic disambiguation and even in the different contexts. This visual observation receives support from the statistical models: For none of the prosodic cues, did the statistical models reveal a main effect of age group. That is, for the use of prosodic cues to mark the internal grouping of coordinates, our data do not provide evidence for a general age-related effect. Second, despite the similarity of the produced prosodic cues, the productions of the older group of speakers are more variable than those of the younger ones, an effect that is evident in larger standard deviations and confidence intervals of the model estimates and the raw data. Increased variability with increased age regarding F0 and durational values is in line with findings of previous studies (Scukanec et al. 1992, 1996, Lortie et al. 2015, Santos et al. 2021, among others).

Regarding our first research question, whether older compared to younger speakers show a more extreme use of F0 range, final lengthening, and pause on Name1 and Name2 to mark the internal grouping, our data do not provide evidence for age-related increases in cue use. In absolute measures, though, older speakers produced longer utterances and longer final vowels on Name1 and Name2 than young speakers, which corresponds to a slower speaking rate since all productions had the same number of syllables. A slower speaking rate is in line with previous findings in the literature (Kemper et al. 1995, Scukanec et al. 1996, Harnsberger et al. 2008, Barnes 2013, Smiljanic & Gilbert 2017, Dimitrova

et al. 2018, Tuomainen & Hazan 2018, Hazan et al. 2019, Tuomainen et al. 2019, 2021). Nevertheless, independent of age, speakers in both age groups marked the internal grouping globally in line with the Proximity/Similarity model (Kentner & Féry 2013) using all three cues investigated: In the brack condition, on Name1, speakers of both age groups produced a smaller F0 range and less final lengthening compared to the nobrack condition. This is considered a weakening of the prosodic boundary indicating the sisterhood of the neighbouring element (i.e., Name2 in this case) by Kentner & Féry (2013). On Name2, this pattern was reversed: In the brack condition, speakers of both age groups increased the F0 range and the lengthening of the final segment compared to the nobrack condition and, additionally, inserted a pause after Name2 in the brack condition. This increase of prosodic cues is considered a strengthening of a prosodic boundary (Kentner & Féry 2013). For none of the prosodic cues was the interaction between age group and condition statistically significant. We, thus, did not find support for age-related more extreme use of disambiguating prosodic cues. Across both age groups, the results of the perception checks confirmed that the internal grouping was produced successfully, as the conditions could reliably be recovered by naïve listeners. Only about 4% of the data in each age group led to misunderstandings. That is, despite the variability in the data, speakers of both age groups produced the disambiguating prosodic cues in such a clear way that listeners could correctly resolve the underlying syntactic structure.

Regarding our second research question, whether young and older speakers differ in adapting their use of prosodic cues when addressing varying interlocutors, our data show substantial similarities across age groups. For several model predictions, the estimated means of the non-baseline contexts within one condition deviate in the same direction from the young baseline context in both age groups (cf. brack in Figures 7 and 8). This also explains why only few interactions of context, condition, and age group revealed statistical significance. Nevertheless, there are slight differences between the age groups regarding their adaptations. We will focus our discussion on statistically significant three-way-interactions of age groups, contexts, and condition, as we are mainly interested in the interplay of all three factors.

The two age groups diverged most strongly when addressing a child as compared to a young interlocutor: On Name2, the older speakers produced larger F0 ranges for the child compared to the young interlocutor in condition brack and smaller F0 ranges along with decreased final lengthening in condition nobrack, thus increasing the difference between conditions when addressing the child. Younger speakers, however, rather slightly decreased the difference between brack and nobrack when addressing the child as they reduced final length-

ening in the brack condition. This enhanced difference between conditions in older speakers can be interpreted as more adaptation to the child interlocutor in older than in younger speakers. Such an enhanced difference between conditions in older compared to younger speakers also holds true for the context with the elderly interlocutor. Here, the older speakers slightly increased the difference between the conditions by means of an increase in final lengthening on Name2 in the brack condition while the young speakers showed the reverse pattern: They decreased the difference in final lengthening between the conditions by increasing final lengthening in nobrack. Interestingly, from the viewpoint of disambiguation, speakers in both age groups produced a stronger distinction between conditions when addressing their peer compared to addressing a non-age-matched interlocutor: young speakers addressing the young interlocutor and older speakers addressing the elderly interlocutor. We are not aware of any similar findings in the literature. Yet, despite being statistically significant, these differences in adaptation between age groups were in fact quite small in absolute terms, and did not affect the disambiguation of coordinates, as revealed by the perception check. Together with the large variability in the productions of the older speakers (cf. larger 95% confidence intervals in the Figures with model predictions than for younger speakers), it is questionable whether the effects in the child and elderly contexts compared to the young context are reproducible in the same manner in future studies.

In the remaining two contexts (non-native interlocutor and speech in noise), our data did not demonstrate evidence for differences between the age groups. Given this and given the fact that any context differences across groups did not impact on disambiguation of coordinates in general, regarding our second research question, our data speak in favour of situational independence in both age groups (Schafer et al. 2000, Kraljic & Brennan 2005, Speer et al. 2011). Models of situational independence assume that disambiguating prosody is realised automatically as part of the production process on the side of the speaker and is therefore largely independent of the presence or absence of a listener, the type of listener, or the situational setting. As such, it seems plausible that disambiguating prosody is also independent of the age of the speaker. Our data add to the literature on the effects of different types of interlocutors and the absence/presence of noise on the use of disambiguating prosodic cues the dimension of speaker age. The findings show that situational independence in production of disambiguating prosody holds for older speakers, too, and that prosody production is a stable automatic part of the production process also in older speakers.

Thus, whereas age has frequently been shown to affect other areas of language production (i.e., word-finding abilities, increased phonetic variability, or altered

acoustic characteristics), it does not seem to have a (listener-relevant) impact on production of prosodic cues in ambiguous structures. This is in line with an observation by Lortie et al. (2015) regarding a more variable voice in older speakers that did not interact with the ability to control fundamental frequency (participants in their study were asked to produce normal, low, and high frequency voice in sustained vowels). In this sense, our study provides evidence that one important part of the prosody-syntax interface is not modulated by age effects: the use of the prosodic cues F0 range, final lengthening, and pause for disambiguation of structurally ambiguous coordinates. Our findings on prosody production in older adults are also of importance in the larger context of investigating linguistic prosody in populations with acquired language and communication disorders resulting from brain lesions (i.e., aphasia or right-hemisphere brain lesions), since participants in these studies are usually older than the typical age groups covered in most studies on healthy prosody processing.

In summary, our data confirm the well-known general age-related changes in absolute durational measures. However, when it comes to the use of tonal and durational prosodic cues to disambiguate the underlying syntactic structure, older speakers modulated duration and F0 range similarly to younger speakers with, if at all, only minimal differences between the age groups of speakers in our sample. The finding of limited adaptation to different interlocutors favours models of situational independence of disambiguating prosody across both age groups and shows that production of disambiguating prosody at the prosody-syntax interface is unaffected by age.

5 Conclusion

In conclusion, young and older speakers in our production study globally marked the internal grouping of coordinated name sequences using F0 range, final lengthening, and pause in a similar way. The modulation of disambiguating prosodic cues seems to be independent of age-related changes in absolute durations. Across both age groups, the use of prosodic cues to resolve the ambiguity in the internal structure of coordinates dominated in comparison to possible prosodic accommodations to the contexts, which we interpret as evidence for situational independence of disambiguating prosody. Prosodic disambiguation thus turns out to be a stable automatic part of the production process, regardless of speaker age.

Acknowledgements

We would like to thank L. Junack for her help with acquisition and pre-processing of the data, as well as for setting up the perception check for the data of the older speakers.

Funding information

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), SFB 1287 project B01 (project no. 317633480).

References

- Barnes, Daniel R. 2013. *Age-related changes to the production of linguistic prosody*. Purdue University. (Doctoral dissertation). https://docs.lib.purdue.edu/open_access_theses/17.
- Bates, Douglas, Reinhold Kliegl, Shravan Vasishth & Harald Baayen. 2015. Parsimonious mixed models. *arXiv: Methodology*. DOI: 10.48550/arXiv.1506.04967.
- Biersack, Sonja, Vera Kempe & Lorna Knapton. 2005. Fine-tuning speech registers: A comparison of the prosodic features of child-directed and foreigner-directed speech. In *Proceedings of InterSpeech*, 2401–2404. DOI: 10.21437/InterSpeech.2005-46.
- Boersma, Paul & David Weenink. 2019. *Praat: Doing phonetics by computer*. <http://www.praat.org/>.
- Burke, Deborah M. & Meredith A. Shafto. 2004. Aging and language production. *Current Directions in Psychological Science* 13(1). 21–24. DOI: 10.1111/j.0963-7214.2004.01301006.x.
- DePaulo, Bella M. & Lerita M. Coleman. 1986. Talking to children, foreigners, and retarded adults. *Journal of Personality and Social Psychology* 51(5). 945–959. DOI: 10.1037/0022-3514.51.5.945.
- Dimitrova, Snezhina, Bistra Andreeva, Christoph Gabriel & Jonas Grünke. 2018. Speaker age effects on prosodic patterns in Bulgarian. In *Proceedings of Speech Prosody*, 709–713. DOI: 10.21437/SpeechProsody.2018-144.
- Dromey, Christopher & Sarah Scott. 2016. The effects of noise on speech movements in young, middle-aged, and older adults. *Speech, Language and Hearing* 19(3). 131–139. DOI: 10.1080/2050571X.2015.1133757.

- Gollrad, Anja, Esther Sommerfeld & Frank Kügler. 2010. Prosodic cue weighting in disambiguation: Case ambiguity in German. In *Proceedings of Speech Prosody*, 165. https://www.isca-speech.org/archive/pdfs/speechprosody_2010/gollrad10_speechprosody.pdf.
- Goy, Huiwen, David N. Fernandes, M. Kathleen Pichora-Fuller & Pascal van Lieshout. 2013. Normative voice data for younger and older adults. *Journal of Voice* 27(5). 545–555. DOI: 10.1016/j.jvoice.2013.03.002.
- Hansen, Marie, Clara Huttenlauch, Carola de Beer, Isabell Wartenburger & Sandra Hanne. 2022. Individual differences in early disambiguation of prosodic grouping. *Language and Speech OnlineFirst*. DOI: 10.1177/00238309221127374.
- Harnsberger, James D., Rahul Shrivastav, William S. Brown Jr, Howard Rothman & Harry Hollien. 2008. Speaking rate and fundamental frequency as speech cues to perceived age. *Journal of Voice* 22(1). 58–69. DOI: 10.1016/j.jvoice.2006.07.004.
- Hazan, Valerie, Outi Tuomainen, Jeesun Kim & Chris Davis. 2019. The effect of visual cues on speech characteristics of older and younger adults in an interactive task. In *Proceedings of the 19th ICPhS*, 815–819. https://www.assta.org/proceedings/ICPhS2019/papers/ICPhS_864.pdf.
- Heister, Julian, Kay-Michael Würzner, Johannes Bubenzer, Edmund Pohl, Thomas Hanneforth, Alexander Geyken & Reinhold Kliegl. 2011. dlexDB: Eine lexikalische Datenbank für die psychologische und linguistische Forschung. In *Psychologische Rundschau*, vol. 62, 10–20. Göttingen: Hogrefe. DOI: 10.1026/0033-3042/a000029.
- Holzgrefe-Lang, Julia, Caroline Wellmann, Caterina Petrone, Romy Råling, Hubert Truckenbrodt, Barbara Höhle & Isabell Wartenburger. 2016. How pitch change and final lengthening cue boundary perception in German: Converging evidence from ERPs and prosodic judgements. *Language, Cognition and Neuroscience* 31(7). 904–920. DOI: 10.1080/23273798.2016.1157195.
- Howell, David C., Marylène Rogier, Vincent Yzerbyt & Yves Bestgen. 1998. *Statistical methods in human sciences*. New York: Wadsworth.
- Huttenlauch, Clara, Carola de Beer, Sandra Hanne & Isabell Wartenburger. 2021. Production of prosodic cues in coordinate name sequences addressing varying interlocutors. *Laboratory Phonology* 12(1). 1–31. DOI: 10.5334/labphon.221.
- Jessen, Michael, Olaf Köster & Stefan Gfroerer. 2003. Effect of increased vocal effort on average and range of fundamental frequency in a sample of 100 German-speaking male subjects. In *Proceedings of the 15th ICPhS*, 1623–1626. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/p15_1623.html.

- Junqua, Jean-Claude. 1996. The influence of acoustics on speech production: A noise-induced stress phenomenon known as the Lombard reflex. *Speech Communication* 20(1). 13–22. DOI: 10.1016/S0167-6393(96)00041-6.
- Kemper, Susan, Patrice Ferrell, Tamara Harden, Andrea Finter-Urczyk & Catherine Billington. 1998. Use of elderspeak by young and older adults to impaired and unimpaired listeners. *Aging, Neuropsychology, and Cognition* 5(1). 43–55. DOI: 10.1076/anec.5.1.43.22.
- Kemper, Susan, Dixie Vandeputte, Karla Rice, Him Cheung & Julia Gubarchuk. 1995. Speech adjustments to aging during a referential communication task. *Journal of Language and Social Psychology* 14(1). 40–59. DOI: 10.1177/0261927X95141003.
- Kentner, Gerrit & Caroline Féry. 2013. A new approach to prosodic grouping. *The Linguistic Review* 30. 277–311. DOI: 10.1515/tlr-2013-0009.
- Knoll, Monja A., Melissa Johnstone & Charlene Blakely. 2015. Can you hear me? Acoustic modifications in speech directed to foreigners and hearing-impaired people. In *Proceedings of InterSpeech*, 2987–2990. DOI: 10.21437/Interspeech.2015-618.
- Knoll, Monja A. & Lisa Scharrer. 2007. Acoustic and affective comparisons of natural and imaginary infant-, foreigner- and adult-directed speech. In *Proceedings of InterSpeech*, 1414–1417. DOI: 10.21437/Interspeech.2007-29.
- Knoll, Monja A., Lisa Scharrer & Alan Costall. 2011. “Look at the shark”: Evaluation of student- and actress-produced standardised sentences of infant- and foreigner-directed speech. *Speech Communication* 53(1). 12–22. DOI: 10.1016/j.specom.2010.08.004.
- Kraljic, Tanya & Susan E. Brennan. 2005. Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology* 50(2). 194–231. DOI: 10.1016/j.cogpsych.2004.08.002.
- Lange, Kristian, Simone Kühn & Elisa Filevich. 2015. “Just Another Tool for Online Studies” (JATOS): An easy solution for setup and management of web servers supporting online studies. *PLoS ONE* 10(6). e0130834. DOI: 10.1371/journal.pone.0130834.
- Lin, Yuri, Jean-Baptiste Michel, Erez Aiden Lieberman, Jon Orwant, Will Brockman & Slav Petrov. 2012. Syntactic annotations for the Google Books Ngram Corpus. In *Proceedings of the 50th annual meeting of the Association for Computational Linguistics*, 169–174. <https://aclanthology.org/P12-3029.pdf>.
- Lombard, Étienne. 1911. Le signe de l’élévation de la voix. *Annales des Maladies de l’Oreille et du Larynx* 37. 101–109.

- Lortie, Catherine L., Mélanie Thibeault, Matthieu J. Guitton & Pascale Tremblay. 2015. Effects of age on the amplitude, frequency and perceived quality of voice. *AGE* 37(6). 1–24. DOI: 10.1007/s11357-015-9854-1.
- Markó, Alexandra & Judit Bóna. 2010. Fundamental frequency patterns: The factors of age and speech type. In *Proceedings of the “Workshop sociophonetics, at the crossroads of speech variation, processing and communication”*, 45–48. http://edizioni.sns.it/wp-content/uploads/2013/12/Sociophonetics_12_MARKO.pdf.
- Mathôt, Sebastiaan, Daniel Schreij & Jan Theeuwes. 2012. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods* 44(2). 314–324. DOI: 10.3758/s13428-011-0168-7.
- Nasreddine, Ziad S., Natalie A. Phillips, Valérie Bédirian, Simon Charbonneau, Victor Whitehead, Isabelle Collin, Jeffrey L. Cummings & Howard Chertkow. 2005. The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society* 53(4). 695–699. DOI: 10.1111/j.1532-5415.2005.53221.x.
- Neurobehavioural Systems. 2018. Version 20.1. <https://www.neurobs.com/>.
- Olusanya, Bolajoko O., Adrian C. Davis & Howard J. Hoffman. 2019. Hearing loss grades and the international classification of functioning, disability and health. *Bulletin of the World Health Organization* 97(10). 725–728. DOI: 10.2471/BLT.19.230367.
- Peters, Benno, Klaus J. Kohler & Thomas Wesener. 2005. Phonetische Merkmale prosodischer Phrasierung in deutscher Spontansprache. In Klaus J. Kohler, Felicitas Kleber & Benno Peters (eds.), *Prosodic structures in German spontaneous speech*, 143–184. Kiel: Arbeitsberichte (AIPUK). https://www.ipds.uni-kiel.de/kjk/pub_exx/aipuk35a/aipuk35a_6.pdf.
- Petrone, Caterina, Hubert Truckenbrodt, Caroline Wellmann, Julia Holzgrefe-Lang, Isabell Wartenburger & Barbara Höhle. 2017. Prosodic boundary cues in German: Evidence from the production and perception of bracketed lists. *Journal of Phonetics* 61. 71–92. DOI: 10.1016/j.wocn.2017.01.002.
- Piazza, Giorgio, Clara D. Martin & Marina Kalashnikova. 2021. The acoustic features and didactic function of foreigner directed speech: A literature review. DOI: 10.31234/osf.io/scnke.
- R Core Team. 2018. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Rodriguez-Cuadrado, Sara, Cristina Baus & Albert Costa. 2018. Foreigner talk through word reduction in native/non-native spoken interactions. *Bilingualism: Language and Cognition* 21(2). 419–426. DOI: 10.1017/S1366728917000402.

- Rojas, Sandra, Elaina Kefalianos & Adam Vogel. 2020. How does our voice change as we age? A systematic review and meta-analysis of acoustic and perceptual voice data from healthy adults over 50 years of age. *Journal of Speech, Language, and Hearing Research* 63(2). 533–551. DOI: 10.1044/2019_jslhr-19-00099.
- Santos, Aline Oliveira, Juliana Godoy, Kelly Silverio & Alcione Brasolotto. 2021. Vocal changes of men and women from different age decades: An analysis from 30 years of age. *Journal of Voice* OnlineFirst. DOI: 10.1016/j.jvoice.2021.06.003.
- Schafer, Amy J., Shari R. Speer, Paul Warren & S. David White. 2000. Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research* 29(2). 169–182. DOI: 10.1023/A:1005192911512.
- Schubö, Fabian & Sabine Zerbian. 2023. The patterns of pre-boundary lengthening in German. In Fabian Schubö, Sabine Zerbian, Sandra Hanne & Isabell Wartenburger (eds.), *Prosodic boundary phenomena*, 1–34. Berlin: Language Science Press. DOI: 10.5281/zenodo.7777526.
- Scukanec, Gail P., Linda Petrosino & Roger D. Colcord. 1996. Age-related differences in acoustical aspects of contrastive stress in women. *Folia Phoniatrica et Logopaedica* 48(5). 231–239. DOI: 10.1159/000266414.
- Scukanec, Gail P., Linda Petrosino & Michael P. Rastatter. 1992. Fundamental frequency variability in elderly women during production of stressed and unstressed words. *Perceptual and Motor Skills* 74(3_suppl). 1091–1095. DOI: 10.2466/pms.1992.74.3c.1091.
- Smiljanic, Rajka & Rachael C. Gilbert. 2017. Acoustics of clear and noise-adapted speech in children, young, and older adults. *Journal of Speech, Language, and Hearing Research* 60(11). 3081–3096. DOI: 10.1044/2017_JSLHR-S-16-0130.
- Smith, Caroline L. 2007. Prosodic accommodation by French speakers to a non-native interlocutor. In *Proceedings of the 16th ICPHS*, 1081–1084. <http://www.icphs2007.de/conference/Papers/1116/1116.pdf>.
- Speer, Shari R., Paul Warren & Amy J. Schafer. 2011. Situationally independent prosodic phrasing. *Laboratory Phonology* 2(1). 35–98. DOI: 10.1515/labphon.2011.002.
- Summers, W. Van, David B. Pisoni, Robert H. Bernacki, Robert I. Pedlow & Michael A. Stokes. 1988. Effects of noise on speech production: Acoustic and perceptual analyses. *The Journal of the Acoustical Society of America* 84(3). 917–928. DOI: 10.1121/1.396660.
- Tauber, Sarah K., Lori E. James & Paula M. Noble. 2010. The effects of age on using prosody to convey meaning and on judging communicative effectiveness. *Psychology and Aging* 25(3). 702–707. DOI: 10.1037/a0019266.

- Thimm, Caja, Ute Rademacher & Lenelis Kruse. 1998. Age stereotypes and patronizing messages: Features of age-adapted speech in technical instructions to the elderly. *Journal of Applied Communication Research* 26(1). 66–82. DOI: 10.1080/00909889809365492.
- Tucker, Benjamin V., Catherine Ford & Stephanie Hedges. 2021. Speech aging: Production and perception. *Wiley Interdisciplinary Reviews: Cognitive Science* 12. e1557. DOI: 10.1002/wcs.1557.
- Tuomainen, Outi & Valerie Hazan. 2018. Investigating clear speech adaptations in spontaneous speech produced in communicative settings. In Mária Gósy & Tekla Etelka Grácsi (eds.), *Challenges in analysis and processing of spontaneous speech*, 9–25. Budapest: Hungarian Academy of Sciences. DOI: 10.18135/CAPS S.9.
- Tuomainen, Outi, Valerie Hazan & Linda Taschenberger. 2019. Speech communication in background noise: Effects of aging. In *Proceedings of the 19th ICPhS*, 805–809. https://www.assta.org/proceedings/ICPhS2019/papers/ICPhS_854.pdf.
- Tuomainen, Outi, Linda Taschenberger, Stuart Rosen & Valerie Hazan. 2021. Speech modifications in interactive speech: Effects of age, sex and noise type. *Philosophical Transactions of the Royal Society B: Biological Sciences* 377(1841). 20200398. DOI: 10.1098/rstb.2020.0398.
- Turk, Alice E., Satsuki Nakai & Mariko Sugahara. 2006. Acoustic segment durations in prosodic research: A practical guide. In Stefan Sudhoff, Denisa Lenerová, Roland Meyer, Sandra Pappert, Petra Augurzky, Ina Mleinek, Nicole Richter & Johannes Schließer (eds.), *Methods in empirical prosody research*, 1–28. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110914641.1.
- Uther, Maria, Monja A. Knoll & Denis Burnham. 2007. Do you speak E-NG-LI-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication* 49(1). 2–7. DOI: 10.1016/j.specom.2006.10.003.
- Wellmann, Caroline, Julia Holzgreife-Lang, Hubert Truckenbrodt, Isabell Wartenburger & Barbara Höhle. 2023. Developmental changes in prosodic boundary cue perception in German-learning infants. In Fabian Schubö, Sabine Zerbian, Sandra Hanne & Isabell Wartenburger (eds.), *Prosodic boundary phenomena*, 119–156. Berlin: Language Science Press. DOI: 10.5281/zenodo.7777532.
- Zollinger, Sue Anne & Henrik Brumm. 2011. The evolution of the Lombard effect: 100 years of psychoacoustic research. *Behaviour* 148(11). 1173–1198. DOI: 10.1163/000579511X605759.