Speech Prosody 2022
23-26 May 2022, Lisbon, Portugal

The effect of intonational rises on serial recall in German

Christine T. Röhr[1], *Michelina Savino*[2,1], *Martine Grice*[1]

[1]I*f*L Phonetik, University of Cologne, Germany
[2]Department of Education, Psychology, Communication, University of Bari, Italy
christine.roehr@uni-koeln.de, michelina.savino@uniba.it, martine.grice@uni-koeln.de

Abstract

This paper uses a serial recall task to investigate the role of rising intonation in the allocation of attentional resources in German. It has been shown for Italian that rising intonation at prosodic boundaries enhances recall of digits in auditorily presented lists. Since resources are usually allocated to prominent items, and since pitch accents are primary encoders of prominence in both languages, we investigate whether an accentual rise leads to better recall than a boundary rise. In a serial recall task on nine-digit sequences in German we compare the effect on working memory of sequences grouped by marking the last item of the two non-final triplets with (i) a high/rising accent followed by an equally high boundary, (ii) a low accent followed by a boundary rise, or (iii) a low/falling accent-boundary sequence, as compared to (iv) ungrouped sequences as controls. Results reveal that items with a rise are recalled more accurately than items without a rise, with no evidence for superior recall of items with accent rises over those with boundary rises. However, boundary rises appear to facilitate recall over a larger domain than accentual rises.

Index Terms: intonation, rises, prominence, attention orienting, serial recall, working memory, German

1. Introduction

In speech processing, prosodic prominence is crucial for directing listeners' attention to the most important parts of the linguistic message. For instance, in West-Germanic languages important words are typically marked by a specific pitch accent [1, 2, 3, 4, 5, 6, 7]. Some evidence also comes from memory tasks involving the recall of prosodically prominent vs. non-prominent words. For instance, in a recognition memory task on English and Korean [8], participants listened to blocks of sentences and at the end of each block, words were presented to them on a screen and listeners were asked whether they had heard the presented word or not in the preceding block. It was found that English participants were more accurate and faster in recalling prosodically prominent than non-prominent words. In another word recognition memory task on American English sentences containing contrastive pitch accentuation, either a simple H* peak accent or a rising L+H* accent [9], it was found that recognition accuracy increased on L+H* accented words. Indications that contrastive rising accents improve word recall were also found for German [10].

The results in [9] and [10] are compatible with findings in other prosody research showing that rising f0 contours are perceived as more prominent than falling ones (see e.g. [11] for German pitch accent types), as rises (and high pitch in general) consume more attentional resources than falls [12] (see also [13] for German). In autosegmental-metrical models of intonation, it is generally assumed that in languages with both pitch accents and boundary tones (i.e. West-Germanic languages including English and German, and also Romance languages including Italian, Spanish and Portuguese), pitch accents are the primary encoders of prominence [14, 15]. This suggests that accentual rises orient listeners' attention more than boundary rises. However, in a serial recall task on auditory nine-digit sequences in Italian [16], it has been shown that rising contours at prosodic boundaries of non-final triplets enhance recall, especially of digits in that position, indicating that boundary tones may also cue prominence.

In this paper, we aim to further explore the functional contribution to prominence of rising pitch in different positions for German. We use a serial recall experiment to compare the effect on recall performance of accentual rises with that of boundary rises. Serial recall tasks require participants to recall a list of items (e.g. digits) in a specific order, usually the order in which they were presented, and are commonly used to assess working memory capacity in research and neuropsychological and psychological settings. Experiments with working memory such as these give us an indication as to how far rising boundary tones have a similar effect to rising accents in the allocation of attentional resources (see [17] for the role of attention and salience in working memory). In a web-based serial recall task (including naturalness ratings) we tested prerecorded nine-digit number sequences that were grouped into triplets by three intonational patterns differing in prosodic prominence. Accordingly, we varied the intonation on the last item in the two non-final triplets in a sequence and tested (i) a *high/rising pitch accent + high boundary* as well as (ii) a *low pitch accent + rising boundary* and (iii) a *low/falling pitch accent + low boundary* contour. An additional ungrouped sequence served as a control condition, which is predicted to achieve the worst recall accuracy (and naturalness evaluation). For the three grouped-by-intonation conditions, we conjecture that if rising boundary tones have a similar effect to rising accents in orienting the listener's attention, in contrast to falling boundary tones, both kinds of rises will lead to an improvement in recall accuracy of the whole sequence as well as of the item bearing the rise. Moreover, we expect a greater beneficial effect on recall performance of accentual rises over boundary rises, as in the former not only the boundary but also the pitch accent is high/rising, and thus more prominent, than in the latter, where only the boundary is rising.

2. Method

2.1. Conditions

As stimuli, we used nine-digit sequences with random ordering of the digits 1 to 9, produced with one of four different intonational patterns. Even though German digits are mostly

 10.21437/SpeechProsody.2022-154

monosyllabic (except for 7), we previously verified in a pilot study that they allow for enough time for the different intonation contours to unfold: 1 *eins* [ʔaɪ̯ns], 2 *zwei* [t͡svaɪ̯], 3 *drei* [draɪ̯], 4 *vier* [fiːɐ̯], 5 *fünf* [fʏnf], 6 *sechs* [zɛks], 7 *sieben* [ˈziːbm̩], 8 *acht* [ʔaxtʰ], 9 *neun* [nɔɪ̯n].

Figure 1 schematically displays the different prosodic conditions: For auditorily presented lists, it is well established that recall is facilitated when they are presented in chunks, separated by pauses, or when intonation is involved in chunking [18, 19, 20]. It has also been found that recall of nine-digit sequences is best improved by grouping in threes [21, 22]. To explore such grouping effects in three grouped-by-intonation conditions, we varied the intonation on the last item in each of the first two (non-final) triplets, i.e. of the items on position 3 and 6 within a sequence. These items were realized with (i) a high/rising accent followed by an equally high boundary (referred to as "accent_RISE": (L+)H* H- according to GToBI [23]), (ii) a low accent followed by a boundary rise ("boundary_RISE": L* H-) and (iii) a low/falling accent-boundary sequence ("boundary_FALL": (H+)L* L-). All three grouped-by-intonation conditions end with a low/falling intonation on position 9 ((H+)L* L-%). A sequence with a neutral, shallow rising-falling intonation (H*) on each item, and without any form of grouping served as a control condition ("ungrouped").

In order to increase naturalness and reduce further influencing factors, we used a similar intonation on the first two items in each triplet for all grouped-by-intonation conditions, in that there is always a downward trend in the f0 contour from the first (H*) to the second (H+!H*) item. The boundary_RISE and accent_RISE conditions correspond to natural list intonation patterns in German (cf. [24, 25]), as high/rising intonation at the end of non-final groups commonly signals continuation or non-finality. In contrast, the boundary_FALL pattern, should be less natural within the list due to its function of signaling finality.

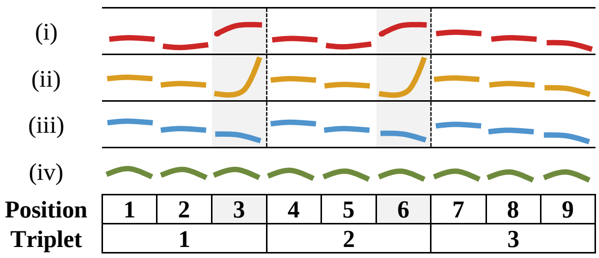

Figure 1: *Schematized pitch contours for the four prosodic conditions. Vertical dashed lines mark intonational group boundaries.*

2.2. Stimuli preparation

The stimuli were constructed according to the procedure reported in [16]: For all digits from 1 to 9, we first produced sequences of the same digit in all nine positions (e.g. for digit *eins* (1): *eins eins eins eins eins eins eins eins eins*) with each of the four prosodic conditions. Thereby we obtained all intonational realizations for each position and prosodic condition for every digit, accounting for downtrends in fundamental frequency (F0) across stretches of natural speech [26]. These sequences were produced as naturally as possible by a trained 37-year-old female native German phonetician and were recorded with a sampling rate of 44100 Hz and 16 bit resolution (mono). No adjustments of the recorded sequences were made, except for an equalization of the sound level. The different digit renditions served as "building blocks" for creating the stimuli for the experimental conditions: All digit renditions were saved as individual audio files and were then concatenated, without pauses, into different nine-digit sequences (containing any digit from 1 to 9) by using Praat [27]. An example of a digit sequence for each of the four prosodic conditions is shown in Figure 2.

We produced 17 sequences for each experimental condition, resulting in a total of 68 stimuli, including four stimuli (one per condition) to be used as sample items. The duration of the stimuli sequences averaged 6.2 seconds (SD = 0.2). The 68 sequences were derived by pseudo-random permutation of the digits 1-9, avoiding two adjacent digits in ascending or descending order within a sequence and avoiding the same digit in an identical position in consecutive sequences within the same condition. In addition, we produced another 14 sequences of digit renditions with neutral intonation (also used for the ungrouped, control condition), for a digit span task conducted prior to the main experiment (see 2.3.).

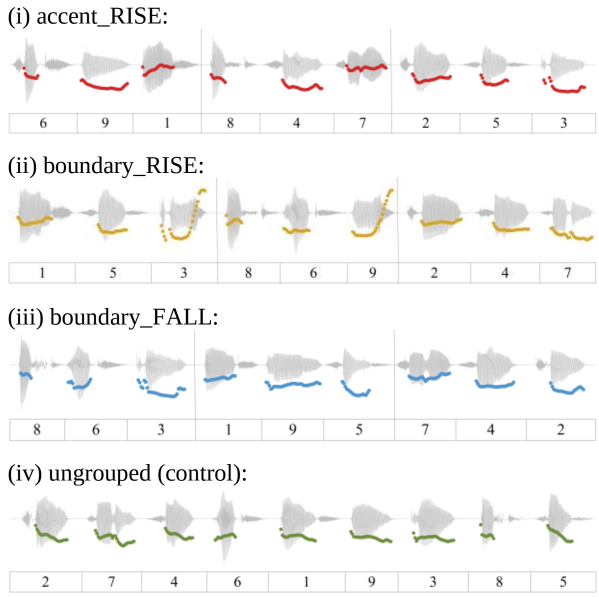

Figure 2: *Speech waveform and F0 contour (100-350 Hz) of a sample sequence for each condition.*

2.3. Procedure

The study was web-based, implemented with the *SoSci Survey* software and made available to participants at www.soscisurvey.de [28]. The participants performed an *immediate serial recall* task: They were instructed to listen to different prerecorded nine-digit sequences (containing digits from 1 to 9) and to recall all nine digits of a sequence in the order in which they were presented immediately after the presentation of the last digit. They entered the digits by clicking each digit on a numeric keypad, with a 3 by 3 grid as e.g. used in cell phones, in the appropriate order. No digit was to be omitted from the input, even in the case of uncertainty. A counter at the top left above the keypad displayed how many digits had already been entered. Participants were able to control when to start a stimulus, but it could only be played once. All stimuli were initiated by a 263 Hz tone of 890 ms, followed by 500 ms of silence before the digit sequence started. The numeric keypad was only displayed after a sequence was completely played. As soon as the participants entered the last digit of a sequence, they could proceed to the next sequence by clicking a button.

The main part of the study consisted of four test blocks, since all 16 stimuli from the same condition were presented as a block (with the different sequences *within* a block always presented in the same order, see 2.2.). To avoid systematic order effects by the exposure to the prosodic conditions, experimental blocks were presented in all possible condition sequences, resulting in 16 different experimental lists which were balanced across participants. At the end of a block, participants were additionally asked to rate the *naturalness* of the previously heard stimuli by answering the question “How did the sequences you heard before sound to you?”. Judgments were given by placing a mark on a visual analogue scale, encoding interval data, with the left pole labelled “unnatural” (= 1) and the right pole labelled “natural” (= 100) (see Fig. 3). The higher the ratings or values, the higher the degree of perceived naturalness. After an experimental block, participants were encouraged to take a short break (1-5 min.).

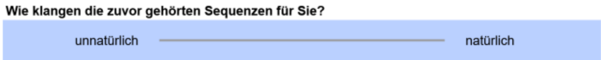

Figure 3: *Example from the naturalness rating task.*

The procedure of the whole study was as follows: At the beginning, participants were asked to use regular stereo headphones and to conduct a volume calibration in order to ensure comparable hearing conditions for all participants. Subsequently, participants were familiarized with the experimental procedure by seeing four sample items (one per prosodic condition). Before starting the main part/task of the study, we also tested the participants’ digit span based on the WAIS-R Digit Span test [29] as control measure. The digit span task, is a serial recall test which requires subjects to repeat series of digits of increasing length and is a widely-used short-term memory test in (neuro-)psychology: After every two-digit sequences of the same length (starting with three digits), we extended the length of the next two sequences by one digit (up to a sequence of nine digits). The digit span corresponds to the length of the last correctly recalled sequence before the participant fails on two consecutive sequences. Finally, the study ended with a question asking for the input modality used (mouse click, touchpad (finger touch) or touchscreen (finger or stylus touch)).

2.4. Participants

Sixty native speakers of German participated in the study. One participant had to be excluded from the analyses for not being raised as a monolingual German speaker, as required. Data from another four participants were discarded due to an exceptionally high performance (i.e. over 90 % recall accuracy in all conditions), since this result is highly unlikely to be gained without additional activity such as writing down the digits whilst listening. The remaining 55 monolingual German participants (26 female, 29 male) are aged between 18 and 47 years (mean age = 29.2 years, SD = 7.8) and their digit span ranges from 4 to 9 with a reasonable mean digit span of 6.8 (SD = 1.2). No participant reported any auditory, visual or neurological impairment. Participants took between 25 to 50 minutes to complete the entire study (mean time = 37.7, SD = 6.2) and most frequently used mouse clicks (69 %) for input, in some cases also the touchpad (24 %) and rarely the touchscreen (7 %).

2.5. Statistical analysis

For the statistical analysis, we performed generalized linear mixed-effects models and linear mixed-effects models by using the *glmer()* and *lmer()* functions from the "lme4" package [30] for R [31]. Generalized linear mixed-effects models were performed on the distribution of correct and false recalls per position/item for the sequences of the main part of the experiment. Items were scored as correct only if they were recalled in the same serial position in which they were presented. We did not consider the first sequence of each condition for the analyses, resulting in a total of 29700 observations (4 conditions * 15 sequences * 9 positions * 55 participants) that entered the analyses. The statistical models, included CONDITION (accent_RISE, boundary_RISE, boundary_FALL, ungrouped (control = reference level)) as the main fixed factor and assume random intercepts and slopes for CONDITION by participants. In addition, we tested the effect of CONDITION on subsets of the data, i.e. for the third position in the first two (non-final) triplets in each sequence (positions 3 and 6; 6600 observations) and for the whole first and second triplet of each sequence (triplet 1 = positions 1, 2, 3 and triplet 2 = positions 4, 5, 6; 9900 observations each). A linear mixed-effects model was performed on the naturalness ratings on the visual analogue scale. The model (220 observations = 4 conditions * 55 participants) included CONDITION as fixed factor and random intercepts by participants. In the following we will only report significant effects (p <= .05).

3. Results

3.1. Recall accuracy

First, results revealed that the recall accuracy (the relative number of digits correctly recalled in the position they were presented) is lowest for the ungrouped, control condition (67.5 %) and highest for the three grouped-by-intonation conditions, which are all quite similar in overall performance (accent_RISE = 78.2 %, boundary_RISE = 80.0 %, boundary_FALL = 77.4 %). A likelihood ratio test revealed that CONDITION had a significant effect on accuracy ($\beta_{acc\_RISE} = .736$, SE = .103, z = 7.176; $\beta_{bound\_RISE} = .771$, SE = .085, z = 9.072, $\beta_{bound\_FALL} = .643$, SE = .099, z = 6.523; $\chi2(3) = 55.177$, $p < .001$). Moreover, pairwise comparisons confirmed that the three grouped-by-intonation conditions significantly differ from the ungrouped condition (with $p < .001$ each), but not from each other.

The Serial Recall Curves in Figure 4 show the accuracy scores for each position in a sequence. The curves reveal a typical U shape which is due to primacy and recency effects [32]: Early items in a sequence are best recalled, followed by the most recent items at the end of the list, with poorer recall of intermediate items. This pattern in particular applies to the ungrouped condition, with the lowest overall recall accuracy. In contrast, the curves of all grouped-by-intonation conditions do not only show an *overall* improvement in recall accuracy but also a *local* improvement on position 6, and for the accent_RISE condition also on position 3. The results on positions 3 and 6 moreover indicate that items produced with a rising pitch, both at the boundary and on the accented syllable, are recalled more accurately than items without this rise (boundary_FALL). A likelihood ratio test on the accuracy on the third position in the two non-final triplets (positions 3 and 6) registered a significant effect of CONDITION ($\beta_{acc\_RISE} = 1.221$, SE = .169, z = 7.242, $\beta_{bound\_RISE} = 1.264$, SE = .158, z = 7.980, β_{bound_FALL} = .861, SE = .157, z = 5.501; $\chi2(3)$ = 50.74, p < .001). Pairwise comparisons revealed significant differences between the three grouped-by-intonation conditions and the ungrouped condition (with p < .001 each), as well as between the accent_RISE and boundary_FALL conditions (p < .05) (boundary_RISE vs. boundary_FALL: p = .056).

Differences throughout the first two (non-final) triplets furthermore suggest superior effects on the *overall* recall performance, when a sequence is grouped by rising boundaries over grouping by rising accents and over grouping by falling boundaries. Likelihood ratio tests on the accuracy on the first and second triplet registered a significant effect of CONDITION (triplet 1: β_{acc_RISE} = .691, SE = .182, z = 3.797, β_{bound_RISE} = .832, SE = .185, z = 4.495, β_{bound_FALL} = .470, SE = .163, z = 2.882; $\chi2(3)$ = 19.989, p < .001; triplet 2: β_{acc_RISE} = .881, SE = .126, z = 6.969, β_{bound_RISE} = 1.093, SE = .110, z = 9.912, β_{bound_FALL} = .725, SE = .123, z = 5.915; $\chi2(3)$ = 60.885, p < .001). Pairwise comparisons revealed significant differences between the three grouped-by-intonation conditions and the ungrouped condition in both non-final triplets (with p < .001 each, except for boundary_FALL in triplet 1 with p < .05), as well as between the boundary_RISE and boundary_FALL conditions in triplet 2 (p < .05).

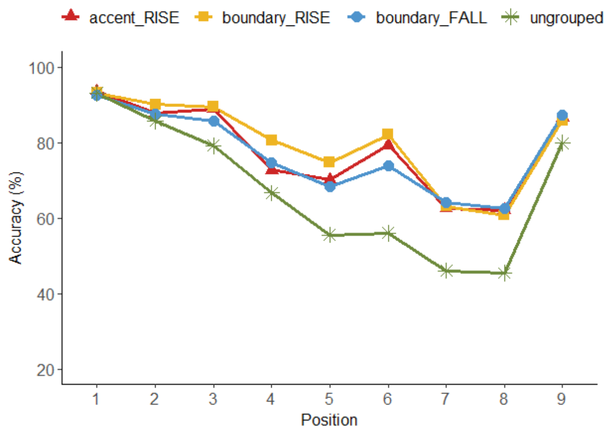

Figure 4: *Serial recall curves by condition.*

3.2. Naturalness

Figure 5 displays the mean values and standard deviations of the naturalness ratings for the different stimuli conditions in combination with a violin plot. The mean ratings reveal that the ungrouped (control) condition was rated lowest in terms of naturalness (mean = 42.4, SD = 28.8), while the three grouped-by-intonation conditions were rated highest, all at a similar level (accent_RISE: mean = 71.9, SD = 25.9, boundary_RISE: mean = 72.4, SD = 27.8, boundary_FALL: mean = 73.6, SD = 22.0). A likelihood ratio test revealed that CONDITION had a significant effect on the ratings (β_{acc_RISE} = 29.455, SE = 3.830, t = 7.690, β_{bound_RISE} = 29.981, SE = 3.830, t = 7.827, β_{bound_FALL} = 31.182, SE = 3.830, t = 8.141; $\chi2(3)$ = 74.073, p < .001). Moreover, pairwise comparisons confirmed that the three grouped-by-intonation conditions significantly differ from the ungrouped condition (with p < .001 each), but not from each other. The violin plots indicating the distribution density of the individual ratings, moreover, suggest slightly different trends between the grouped-by-intonation conditions. What is evident from these plots is that the boundary_FALL condition is *not* less natural than the two rising conditions, counter to our expectations.

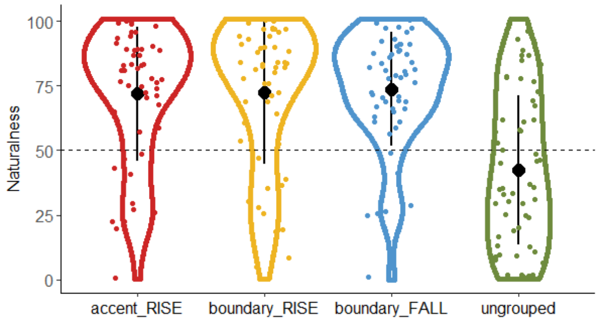

Figure 5: *Violin plots of naturalness ratings (0=unnatural, 100=natural) by condition. Mean values and standard deviations are indicated by black dots and vertical lines.*

4. Discussion and conclusions

In an immediate serial recall task we investigated the role of rising accents and rising boundary tones in the allocation of attentional resources. Results confirmed that intonational grouping of nine-digit sequences into triplets with rising pitch accents, rising boundary tones and falling boundary tones led to better overall recall performances than neutral intonation without grouping. Moreover, statistical analyses revealed that both accentual and boundary rises led to better recall of the items and domains they are associated with, than falling boundaries. We did not find a beneficial effect of accent rises over boundary rises, or vice versa, but it appears that boundary rises have a more global effect on recall performance, facilitating recall of a whole triplet, while accentual rises have a more local effect, with greater improvement on the accented item. The global effect of boundary H tones is unsurprising, as in autosegmental-metrical phonology it is analyzed as a tone associated with the whole phrase (in this case the three-digit sequence), and could thus be facilitating recall of all items in that phrase, as opposed to an accentual H tone, that may have a more local effect, reflecting its association with a single digit.

Naturalness ratings indicate that sequences with intonational grouping are perceived as more natural than ungrouped sequences. Sequences with falling boundaries received high naturalness ratings, which was unexpected but might explain their comparatively high accuracy scores.

In general, our results provide further evidence that intonational grouping improves recall ability. Furthermore, intonational rises led to better recall performances than falls, indicating that they direct more attention towards an item or even adjacent items. If the effect on accuracy in recall can be related to prominence, then boundary rises appear to be lending prominence to the whole domain of which they are a boundary (the triplet), whilst accentual rises, as expected, tend to have a more local effect on the digit which is accented.

5. Acknowledgements

This research has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 281511265 – SFB 1252 “Prominence in Language” (project A01) at the University of Cologne.

6. References

[1] J. B. Pierrehumbert and J. Hirschberg, "The Meaning of Intonational Contours in the Interpretation of Discourse," in P. R. Cohen, J. Morgan and M. E. Pollak (eds.), *Intentions in Communication*, 1990, pp. 271–311. Cambridge: MIT Press.

[2] K. Kohler, "Terminal Intonation Patterns in Single-Accent Utterances of German: Phonetics, Phonology and Semantics," *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK)*, vol. 25, pp. 115–185, 1991.

[3] K. Ito, S. Speer, and M. Beckman, "Informational Status and Pitch Accent Distribution in Spontaneous Dialogues in English," *Proceedings of the 2nd International Conference on Speech Prosody*, pp. 279-282, 2004.

[4] S. Baumann, J. Becker, M. Grice, and D. Mücke, "Tonal and Articulatory Marking of Focus in German," in *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)*, pp. 1029–1032, 2007.

[5] A. Chen, E. Den Os, and J. P. De Ruiter, "Pitch Accent Type Matters for Online Processing of Information Status: Evidence from Natural and Synthetic Speech," *The Linguistic Review*, vol. 24, no. 2–3, pp. 317–344, 2007.

[6] C. T. Röhr and S. Baumann, "Prosodic Marking of Information Status in German," *Proceedings of the 5th International Conference on Speech Prosody*, vol. 100019, pp. 1–4, 2010.

[7] M. Grice, S. Ritter, H. Niemann, and T. B. Roettger, "Integrating the Discreteness and Continuity of Intonational Categories," *Journal of Phonetics*, vol. 64, pp. 90–107, 2017.

[8] H. Kember, J. Choi, J. Yu, and A. Cutler, "The Processing of Linguistic Prominence," *Language and Speech*, vol. 64, no. 2, pp. 413-436, 2019.

[9] S. H. Fraundorf, D. G. Watson, and A. S. Benjamin, "Recognition Memory Reveals Just How CONTRASTIVE Contrastive Accent Really Is," *Journal of Memory and Language*, vol. 63, no. 4, pp. 367-386, 2010.

[10] X. Koch and K. Spalek, "Contrastive Intonation Effects on Word Recall for Information-structural Alternatives Across the Sexes," *Memory and Cognition*, vol. 49, pp. 1312-1333, 2021.

[11] S. Baumann and C. T. Röhr, "The Perceptual Prominence of Pitch Accent Types in German," *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*, vol. 298, pp. 1–5, 2015.

[12] C.-H. Hsu, J. P. Evans, and C.-Y. Lee, "Brain Responses to Spoken F0 Changes: Is H Special?," *Journal of Phonetics*, vol. 51, pp. 82-92, 2015

[13] C. T. Röhr, I. Brilmayer, S. Baumann, M. Grice, and P. B. Schumacher, "Signal-Driven and Expectation-Driven Processing of Accent Types," Language, Cognition & Neuroscience, vol. 36, no. 1, pp. 33-59, 2020.

[14] D. R. Ladd, *Intonational phonology*. Cambridge: Cambridge University Press, 2008.

[15] A. Arvaniti, "The Autosegmental-Metrical Theory of Intonational Phonology," in C. Gussenhoven and A. Chen (eds.), *The Oxford Handbook of Language Prosody*, 2020, pp. 1-20. Oxford: Oxford University Press.

[16] M. Savino, B. Winter, A. Bosco, and M. Grice, "Intonation Does Aid Serial Recall After All," *Psychonomic Bullettin and Review*, vol. 27, pp. 366-372, 2020.

[17] K. Oberauer, "Working Memory and Attention. A Conceptual Analysis and Review," *Journal of Cognition*, vol. 2, no. 1, pp. 1-23, 2019.

[18] C. Frankish, "Modality-specific Grouping Effect in Short-term Memory," *Journal of Memory and Language*, vol. 24, no. 2, pp. 200-209, 1985.

[19] C. Frankish, "Intonation and Auditory Grouping in Immediate Serial Recall," *Applied Cognitive Psychology*, vol. 9, no. 7, pp. 5-22, 1995.

[20] S. Saito, "Effects of Articulatory Suppression on Immediate Serial Recall of Temporarily Grouped and Intonated Lists," *Psychologia*, vol. 41, no. 2, pp. 95-101, 1998.

[21] W. A. Wickelgren, "Size of Rehearsal Group and Short-term Memory," *Journal of Experimental Psychology*, vol. 68, pp. 413-419, 1964.

[22] J. Ryan, "Temporal Grouping, Rehearsal and Short-Term Memory," *Quarterly Journal of Experimental Psychology*, vol. 21, pp. 148–155, 1969.

[23] M. Grice, S. Baumann, and R. Benzmüller, "German Intonation in Autosegmental-Metrical Phonology," in S.-A. Jun (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*, 2005, pp. 55–83. Oxford: Oxford University Press.

[24] S. Baumann and J. Trouvain, "On the Prosody of German Telephone Numbers," *Proceedings 7th Conference on Speech Communication and Technology*, pp. 557-560, 2001.

[25] J. Peters, "Phonological and Semantic Aspects of German Intonation," *Linguistik Online*, vol. 88, no. 1, 2018.

[26] D. R. Ladd, "Declination: A Review and Some Hypotheses," *Phonology*, vol. 1, pp. 53-74, 1984.

[27] P. Boersma and D. Weenink, Praat: Doing Phonetics by Computer [Computer Program], Version 6.1.38, retrieved January 2021 from https://www.praat.org/.

[28] D. J. Leiner, SoSci Survey (Version 3.1.06) [Computer Software], 2019, available at https://www.soscisurvey.de/.

[29] D. Wechsler, *Manual for WAIS-R*. New York, NY: Psychological, 1987.

[30] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting Linear Mixed-Effects Models Using lme4," *Journal of Statistical Software*, vol. 67, no. 1-48, 2015.

[31] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2021, https://www.R-project.org/.

[32] H. Ebbinghaus, *On Memory: A Contribution to Experimental Psychology*. New York: Teachers College, 1913.