COMPENSATION STRATEGIES IN NON-NATIVE ENGLISH AND GERMAN

Katharina Zahner[1] *and Jenny Yu*[2,3]

[1]University of Konstanz, Germany [2]The MARCS Institute, Western Sydney University, Australia
[3]ARC Centre of Excellence for the Dynamics of Language, Australia
katharina.zahner@uni-konstanz.de, jenny.yu@westernsydney.edu.au

ABSTRACT

English and German speakers differ in how they realise pitch contours when there is limited time for voicing. Replicating the pattern found by Grabe [1] for Southern British English vs. Northern German, Yu and Zahner [2] recently showed that Australian English speakers compressed rises and falls in English, while Southern German speakers compressed rises but truncated falls in German. In the present follow up study, we investigate whether speakers also transfer compensation strategies to their L2 by examining non-native productions of English and German contours by speakers in [2]. Results suggest evidence of both transfer and adjustment: Southern German speakers compressed rises and truncated falls in English as they did in German, while Australian English speakers shifted from compression to truncation when producing German falls. We discuss factors that may account for this asymmetrical behaviour in prosodic transfer and implications of these findings for language learning.

Keywords: compensation strategies, L2 production

1. INTRODUCTION

When a spoken word contains limited time for voicing, it can be hard to realise complex intonational contours. Truncation and compression are two strategies speakers can use to “compensate” and cope with this problem. In truncation, the pitch slope remains unchanged and the f0 movement simply stops when voiced material runs out. Compression, by contrast, involves increasing the speed of the movement to accommodate the full contour within a shorter time span (see [3, 4] for other strategies).

Speakers of different languages prefer different strategies [1]. Recently, [2] found a difference for Australian English and Southern German: Participants produced surnames with progressively less voicing (e.g., *Sheafer, Sheaf, Shift*) in question versus statement contexts, requiring rising and falling contours, respectively. English speakers compressed pitch movements when segmental material decreased for both rises and falls. German speakers also compressed rises, but truncated falls.

This paper now reports on the non-native (L2) productions of English and German contours by participants in [2] to investigate whether the (language-specific) use of compression and truncation transfers to non-native speech. Research in this area is relevant as it may provide insights into determining L2 proficiency or help non-native speakers find ways to reduce foreign accent and increase intelligibility (cf. [5] on the interaction between prosody and segments in the perception of foreign accented speech).

To date, research on compensation strategies in L2 productions has been scarce. One study by He, et al. [6] tested Dutch native speakers and Mandarin Chinese L2 speakers of Dutch (with different proficiency levels as judged by experts) on their intonational compensation strategies in rises, falls, and fall-rises. Mandarin Chinese speakers with higher Dutch proficiency were found to realise contours more similarly to native Dutch speakers, i.e., they compressed both rises and falls, while Chinese speakers with lower Dutch proficiency did not show this pattern. For fall-rises, neither subgroup reproduced the native pattern, which the authors attribute to the fact that this contour does not exist in Mandarin. Hence, language proficiency might modulate the use of compensation strategy, as well as native language constraints.

There is also evidence that speakers’ phonetic implementation of pitch accent types in their L1 transfers to L2 speech. Atterer and Ladd [7], for instance, found that German speakers aligned prenuclear accentual rises significantly later than English speakers and these native alignment patterns transferred to productions in English; see also [8].

Against this background, we derive the following predictions: If the use of compression and truncation similarly transfers, then speakers should prefer the same strategies in native and non-native productions. Specifically, for Australian English speakers producing L2 German, we should find evidence of compression in both rises and falls to reflect the strategies found in their native English productions. Similarly, for Southern Germans, we should observe compression in their English rises and truncation in falls. If, as suggested by [6], higher language proficiency may foster more native-like productions, then we may observe the same patterns as in the target language (note that we explicitly address proficient speakers in L2, i.e., self-reported level above B2).

2. METHOD

2.1. Participants

Twelve native Australian English speakers of German (5 females, M_{age} = 33.86 years, SD = 11.65 years) and twelve native German speakers of English (9 females, M_{age} = 25.45 years, SD = 3.98 years) participated for a small payment. Native English speakers were recruited in Sydney and native German speakers were recruited in Tübingen (n = 4) and Konstanz (n = 8). On average, English speakers' self-rated L2 proficiency (1-7 scale) was 4.83 (SD = 1.24), German speakers' self-rated L2 proficiency was 5.27 (SD = 0.84). A non-parametric Wilcoxon test showed that this difference was not significant (W = 48, p = 0.17).

2.2. Stimuli

Table 1 shows the test items: 12 equivalent surnames for English and German which formed four sets of three name types. For each name type, there were three conditions, long (disyllabic), mid (monosyllabic with a long vowel), and short (monosyllabic with a short vowel).

Table 1: Test items in English and German (italicized).

Continuum	Step 3: Long	Step 2: Mid	Step 1: Short
sh	Sheafer	Sheaf	Shift
	Schiefer	*Schief*	*Schiff*
s	Seefer	Seef	Siff
	Siefer	*Sief*	*Siff*
g	Geesser	Geese	Giss
	Gieser	*Gies*	*Giss*
k	Keesser	Keese	Kiss
	Kieser	*Kies*	*Kiss*

Test items occurred in syntactically similar carrier phrases in two lists, in a polar question, designed to elicit a nuclear rising contour on the test word, and in a declarative statement for a nuclear falling contour (see (1) and (2)). The test item, in phrase-final position, was followed by a polysyllabic appositional phrase, which is expected to be realised with the same contour as the test item. This served as a control to determine the underlying intonational specification of a test word in case of truncation. Each carrier also had a short preceding introductory paragraph to provide context (e.g., *Anna and Peter are watching TV. A photograph of this week's National Lottery winner appears. Anna says: ''Look, Peter!''*). A further 28 filler stimuli were created for each language: 14 declaratives and 14 polar questions with surnames in various positions.

(1) Carrier phrases for falls (test items are bolded):
English: It's Mr. **Sheafer**! Our new neighbour!
German: Das ist Herr **Schiefer**! Unser neuer Nachbar!

(2) Carrier phrases for rises:
English: Isn't that Mr. **Sheafer**? Our new neighbour?
German: Ist das nicht Herr **Schiefer**? Unser neuer Nachbar?

2.3. Procedure

Subjects were recorded individually in a sound-attenuated booth (both languages were tested in one session, L1 (see [2]) and L2, but only L2 is reported here). The experiment was controlled using *Presentation* software [9]. To reduce confusion, subjects completed all items in one language before recording items in the other. Between languages, participants read a passage to adjust to speaking in the other language. Items were blocked by contour for each language, such that subjects produced all statements or questions before producing the other type. Language and contour order were counter-balanced and randomly assigned to subjects.

Each trial started with the presentation of an introductory paragraph on screen which subjects were instructed to read silently for context. Subjects then read the corresponding carrier phrase out loud. To prevent subjects from ignoring contextual information, the carrier phrase only appeared underneath the introductory paragraph after three seconds. In addition, participants were asked to speak naturally and not exaggerate their speech. After producing the sentence, participants pressed a button to progress onto the next item.

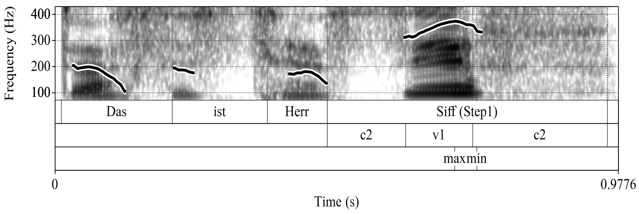
Each block began with three practice trials to familiarise the subject with producing items and for the experimenter to provide feedback e.g., if subjects exaggerated or spoke unclearly. 26 test trials (12 critical trials, 14 filler trials) followed in an individually randomized order for each of the four blocks (104 test trials in total).

3. ACOUSTIC ANALYSIS AND DATA TREATMENT

34 L2 English and 51 L2 German productions were excluded due to errors in pronunciation, missing data, or wrong contour (a fall produced in a statement context or a rise in a question context). The final set consisted of 491 productions (254 L2 English, 237 L2 German).

Word and segment boundaries were annotated following standard segmentation criteria (see [10]). F0 minima and f0 maxima were annotated for each nuclear contour (fall or rise). In monosyllabic words, both min and max occurred in the vowel of the test word, in disyllabic words in different vowels.

Figure 1: Spectrogram and f0 contour with annotation: German "Siff" (step 1), fall, produced by an English speaker. Tier 1 shows the words in German, Tier 2, segments of target word *Siff,* Tier 3 the f0 min and f0 max of the accentual movement.



To analyse effects of voicing duration on f0 movement, "rate of f0 change" (RoCh) was calculated by dividing f0 excursion (f0 maximum minus f0 minimum in semitones, st) of a fall or rise by movement duration. We used semitones instead of Hz to account for gender differences in the data. Comparing across the three continuum steps (from longest to shortest), a RoCh increase is termed compression, and a stable rate of change, truncation. The f0 excursion (in st) was also examined to see whether RoCh modifications go together with a reduction of the f0 range of the accentual movement.

Linear mixed effects regression models in R [11] were fit for the dependent variables (RoCh, f0 excursion). In each model, *language* (English vs. German, spoken by L2 speakers), *contour* (falling vs. rising), and *step* (1, 2, 3) were modelled as fixed factors and *subject* and *word type* (sh-, s-, g-, and k-continuum depending on the language) as crossed random factors [cf. 12, 13]. Random slopes for the fixed factors were added and retained if the model fit improved [14, 15] (LogLikelihood). Planned group comparisons were conducted using the *emmeans* function [16], with Tukey-adjusted p-values.

4. RESULTS AND DISCUSSION

Duration analysis. F0 duration (absolute time difference between f0 min and f0 max) was analysed to check whether the duration manipulation provided successively less sonorant material between steps. For both L2 English and L2 German, the duration of voiced material decreased significantly with each step from long to short, for both falls and rises (individual comparisons, all $p<.0001$), indicating the duration manipulation was successful.

F0 analyses. An increased RoCh and stable f0 excursion on shorter words is interpreted as compression, while a stable RoCh and smaller f0 excursions are evidence of truncation. Figure 2 shows RoCh (st/ms) for L2 falls and rises and Figure 3 illustrates the average f0 excursion (st). The data show that RoCh increases for L2 rises in both languages. L2 English and German falls, however, do not show a step-wise increase (see Figure 2). In Figure 3, there appears to be a general decrease in f0 excursion from "long" to "short" in the L2 falls of both languages while in comparison, f0 excursion in L2 rises appears more stable across steps.

Figure 2. Rate of Change (RoCh) in st/ms for falls (left) and rises (right), split by language (non-native).

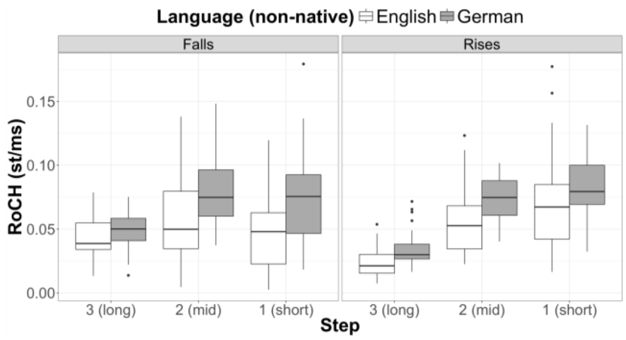
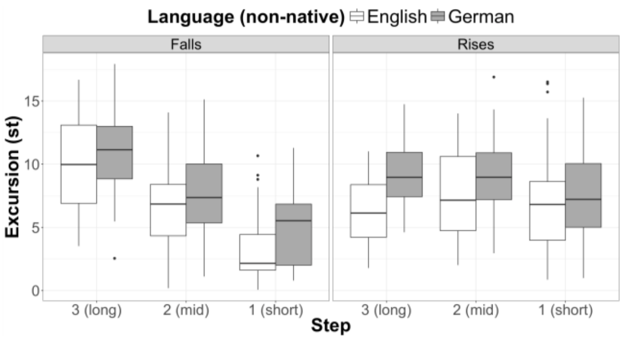


Figure 3. Average f0 excursion in st for falls (left) and rises (right), split by language (non-native).



The **statistical analysis** showed that, for **RoCh**, there was no interaction between *contour* and *language* ($p = .11$), suggesting that rising and falling contours were not realised differently between languages. German and English were therefore pooled. Rises and falls, however, were analysed separately as there was an interaction between *contour* and *step* ($p<.0001$). **For L2 rises**, there was a main effect of *step*, reflecting the increase in RoCh from the longest to the shortest word, i.e., compression. RoCh differed between all individual steps (all $p<.01$). **For L2 falls**, there was also an effect of *step* ($p<.001$). This was mainly driven by the difference between step 2 and step 3 ($p<.001$, see Figure 2), which suggests compression between these steps. However, there was no difference between step 1 and 2 ($p=.20$) and, importantly, no difference between the shortest and the longest word, i.e., step 1 and 3 ($p=.09$), indicating a trend for truncation overall.

For **f0 excursion**, there was a three-way interaction between *step*, *language*, and *contour* (p=.012). **For rises**, step affected f0 excursion differently in L2 German and English (p=.042): **For L2 German rises**, the differences between step 1 vs. 2 and step 2 vs. 3 were not significant (p>0.28), but the f0 excursion was smaller in step 1 than in step 3 (p=.027), indicating that compression was accompanied by some reduction of f0 range. **For L2 English rises,** however, there was no effect of step (all individual comparisons p>.26), corroborating compression. **For falls**, there was no interaction between *language* and *step* (p=.88), indicating that step did not affect f0 excursion differently in L2 German and English. Subsequently, the interaction term was removed from the model. Importantly, there was a main effect of step (p<.001): F0 excursion significantly decreased for each step of the continuum (all ps<0.001). These results further support the use of truncation in L2 English and German falls as indicated from the RoCh analyses.

Analyses found trends for English and German speakers compressing rises and truncating falls. However, results between individual steps were less conclusive, especially for L2 falls. More precisely, **RoCh analysis** for L2 falls did not point conclusively to either compression or truncation. We therefore conducted a post-hoc combined **analysis of RoCh and f0 excursion** for L2 falls (where we multiplied the two measures and used the product as the dependent variable) to confirm that *step* did not affect the combined measure differently in L2 German and English (p>.85), i.e., speakers used the same strategy. To further corroborate the difference in strategy in English speakers when speaking English and German, respectively, we pooled L1 English falls from [2] and L2 German falls (spoken by the same English speakers, present data). Results revealed an interaction between *language* and *contour* (p=.05), supporting the claim that English speakers use different compensation strategies for L1 and L2 falls.

Therefore, results indicate both English and German speakers compress rises and truncate falls in their non-native productions. For German speakers, this pattern of results reflects the strategies found in their native productions (see [2]), suggesting that German speakers transferred their preferred strategy from L1 to L2. English speakers also showed transfer in rises, compressing pitch movements when producing both English and German. However, English speakers tended to truncate falls in German. This is unlike their preference for compression in their L1 (see [2]), and instead resembles the behaviour of native German speakers. Therefore, results suggest evidence of both transfer and adjustment of compensation strategies from native to non-native productions.

This presents an unexpected asymmetric pattern of behaviour between English and German speakers. In [6], speakers with higher L2 proficiency produced realisations more similarly to native speakers. Yet, in our study, we specifically tested proficient L2 speakers. In addition, self-assessed L2 proficiency ratings were not significantly different between German and English groups. Therefore, if anything, both groups should have been expected to adjust their compensation strategies. Articulatory reasons could also explain the asymmetric prosodic transfer. Physiologically, compression requires more effort than truncation as speakers need to increase pitch rate to accommodate the contour. In rising movements, articulatory effort is high in all cases by an increase in laryngeal tension [16] (to counteract declination processes [17]). Therefore, compression may be less effortful in rises (where effort is high anyway), but more difficult to achieve in falls (where an additional effort needs to be made). This could explain a preference for non-native speakers to truncate falls.

5. CONCLUSION

Overall findings suggested English and German speakers compressed rises and truncated falls in their L2, providing evidence that compensation strategies can be both transferred or adjusted from native to non-native productions. Hence, our results paint a complex relationship between transfer and L2 prosody. Clearly, more research is needed to understand the underlying factors underlying the use of compensation strategies in L2. Further investigation will explore patterns between steps (including e.g., durational points and number of syllables), other measures for capturing compensation strategies, and speaker-specific productions to address individual differences. Future studies should also include other language combinations, particularly involving pairings with an asymmetric behaviour in the L1. Continued research may eventually help non-native speakers use compensation strategies more appropriately to reduce foreign-accentedness.

6. AKNOWLEDGEMENTS

We thank Bettina Braun and Anne Cutler for their feedback and guidance. This work was funded by the ARC Centre of Excellence for the Dynamics of Language (project ID: CE140100041) and an Australia-Germany Joint Research Cooperation Scheme award to A. Weber (for SpeechNet BaWü) and A. Cutler.

7. REFERENCES

[1] Grabe, E., "Pitch accent realization in English and German," *Journal of Phonetics,* vol. 26, pp. 129-143, 1998.

[2] Yu, J. and Zahner, K., "Truncation and compression in Southern German and Australian English," *Proceedings of the 19th Annual Conference of the International Speech Communication Association (Interspeech)*, Hyderabad, India, 2018, pp. 1833-1837.

[3] Rathcke, T. V., "How truncating are ‘truncating languages'? Evidence from Russian and German," *Phonetica*, vol. 73, pp. 194-228, 2017.

[4] Hanssen, J., *Regional variation in the realization of intonation contours in the Netherlands*. Utrecht, the Netherlands: LOT, 2017.

[5] Ulbrich, C. and Mennen, I., "When prosody kicks in: The intricate interplay between segments and prosody in perceptions of foreign accent," *International Journal of Bilingualism,* vol. 20, pp. 522-549, 2016.

[6] He, X., Hanssen, J., J., v. H. V., and Gussenhoven, C., "Mandarin-accented fall, rise and fall-rise f0 contours in Dutch," in *Proceedings of the 6th International Conference on Speech Prosody*, Shanghai, China, 2012, pp. 358-361.

[7] Atterer, M. and Ladd, D. R., "On the phonetics and phonology of "segmental anchoring" of F0: Evidence from German," *Journal of Phonetics,* vol. 32, pp. 177-197, 2004.

[8] Gut, U., *Non-native speech. A corpus-based analysis of phonological and phonetic properties of L2 English and German.* Frankfurt: Peter Lang, 2009.

[9] Presentation, N. S., ed. Berkeley, CA, 2000.

[10] Turk, A. E., Nakai, S., and Sugahara, M., "Acoustic segment durations in prosodic research: A practical guide," in *Methods in Empirical Prosody Research*, Sudhoff, S., Lenertová, D., Meyer, R., Pappert, S., Augurzky, P., Mleinek, I., *et al.*, Eds., Berlin, New York: Walter de Gruyter, 2006.

[11] R Development Core Team, R., "R: A language and environment for statistical computing," Vienna: R Foundation for Statistical Computing, 2015.

[12] Baayen, R. H., *Analyzing Linguistic Data. A Practical Introduction to Statistics*. Cambridge: Cambridge University Press, 2008.

[13] Baayen, R. H., Davidson, D. J., and Bates, D. M., "Mixed-effects modeling with crossed random effects for subjects and items," *Journal of Memory and Language*, vol. 59, pp. 390-412, 2008.

[14] Bates, D. M., Kliegl, R., Vasishth, S., and Baayen, H. R., "Parsimonious mixed models," *arXhiv preprint,* vol. arXiv:1506.04967, 2015.

[15] Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H. R., and Bates, D. M., "Balancing type 1 error and power in linear mixed models," *Journal of Memory and Language*, vol. 94, pp. 305-315, 2017.

[16] Lenth, R., "emmeans: Estimated Marginal Means, aka Least-Squares Means, R package version 1.2.3.https://CRAN.R-project.org/package=emmeans," 2018.