ICPhS 2027
August 9 - 13, 2027 | Victoria, Canada

UNDER REVIEW · submitted to ICPhS 2027

EVALUATING SWISS STANDARD GERMAN: A MATCHED-GUISE STUDY

XXX instead of the name(s) of the author(s)
Please write XXX instead of the affiliation(s)
Please write XXX instead of the email address(es)

ABSTRACT

Spoken Swiss Standard German ranges from strongly Swiss to nearly German. Which point listeners prefer is unknown: earlier work varied the situation, not the pronunciation. We varied pronunciation alone. One speaker read one passage as Swiss-near, Mixed and German-near, varying six segmental variables, everything else constant. Each of 98 Swiss listeners heard one guise and rated the speaker on competence (status), likability and attractiveness (solidarity). Mixed scored highest on all three, exceeding German-near by 1.65 rating points on solidarity but only 0.71 on competence. German-near gained no status for the solidarity it lost, and listener background made no difference. Approaching the dominant standard of a pluricentric language, then, costs a speaker of the smaller variety belonging without buying status.
Keywords: Swiss Standard German; matched guise; language attitudes; status and solidarity.

1. INTRODUCTION

German-speaking Switzerland is diglossic: dialect carries everyday interaction, the spoken standard schooling, most broadcasting and formal or non-local encounters [1], [2], [3]. Swiss Standard German has no codified pronunciation norm of its own, only an unwritten one [4]; we call the codified pronunciation the German norm and departures from it Swiss. Hove [5] calls the unwritten norm the *schweizerhochdeutsche Sprachkonvention* ('Swiss Standard German linguistic convention'), a tacit agreement on which variants a Swiss speaker may use, penalizing departure either way: too German sounds conceited, too dialectal uneducated. Actors and broadcasters are criticized both ways, for sounding too Swiss [2], [3] and too German [5]. If both poles are sanctioned, the best-rated point should lie between them, but where, and on which dimensions, is unmeasured.

Which variables make a speaker sound Swiss or German? Guntern [6] groups them by how uniformly Swiss speakers depart from the norm, using the corpora of [5] and [7]. On one variable, almost all speakers depart alike: prevocalic <s> is [z] in the norm but [s] in Switzerland [7]. A second, <ä>, is a stereotype: the over-open [æ] (as opposed to the norm [ɛ]) is obligatory in imitations of "bad" Swiss Standard German, yet Swiss speakers use it only rarely in everyday speech and not at all when reading aloud [5], [7]. On three consonantal variables, speakers go either way: <ch> after a front vowel is [ç] in the German norm but often [x, χ] in Swiss Standard German; <k> is [kʰ] but often affricated [k͡x]; <r> is [ʁ] before a vowel and vocalized after one, but most often apical [r] in both. In numbers: <r> stays consonantal in some 90% of the positions where the norm vocalizes, while <k> is affricated in 35% of everyday tokens and 2% of read ones [7]. A sixth variable, <chs> as in *sechs* ('six'), is realized as [ks] in the norm, [xs] in Switzerland [7]. Listeners appear to hear the three consonantal variables most readily as Swiss [6], and <r> is what they name most often when asked to locate a speaker within Switzerland [8].

Evaluation is less settled. Asked outright, Hove's [5] respondents rated German standard "better" but Swiss standard more likable; two thirds nonetheless found a Swiss speaking like a German *unsympathisch* ('unlikable'). Guntern [2] took two equally Swiss-sounding recordings: a Swiss detective among German colleagues in a TV crime drama, and a voice-over reading a love letter in a cinema ad. Her 31 listeners rated the detective inappropriate and incompetent, the letter the opposite. She reads this as listeners having coexisting *Hochdeutsch-Ideale* ('ideals of standard German'): the same pronunciation fits one situation and not the other. But the two scenarios also differ in speaker, sex, text and genre: what she varied was the situation, not the pronunciation. We know of no study varying only the pronunciation.

We therefore had one speaker read a passage in three guises, Swiss-near, Mixed and German-near, holding text and situation constant, and asked Swiss listeners to rate him. Three hypotheses follow:

- H1. Ratings peak between the poles, since the convention penalizes departure either way [5].
- H2. Guise affects solidarity (likability, attractiveness) more than status (competence) [9], the same line as Hove's split [5].
- H3. The German-near guise is not rated above Swiss-near on competence, even though

Under review. Submitted to ICPhS 2027. Please do not cite without permission.

Hove's respondents called the German standard “better” [5].

2. METHODS

2.1. The experiment

Data come from a web-based matched-guise experiment [10], using SoSci Survey. Swiss German listeners were recruited by WhatsApp and e-mail and supplied metadata: age, gender, highest completed qualification and canton of upbringing. In a between-subjects design, each then heard one recording of a man reading a short standard German passage and rated him on three six-point scales – competence, likability and attractiveness – from 1 (not at all) to 6 (very).

2.2. Listeners

We analyze the 98 listeners who completed the questionnaire, all raised in Switzerland. The sample skews female (62 women, 63.3%; 36 men) and young (M = 27.3, SD = 11.8; median 24; range 15-67), with 76% under 30. Highest qualification was compulsory schooling or a vocational apprenticeship for 45, an intermediate qualification for 21 and a university degree for 30; two gave "other". Coverage is dominated by Bern (74 of 98); eastern Switzerland is barely represented (n = 2). The three guises drew 34, 32 and 32 listeners with comparable sex ratios (20, 20 and 22 women), but unequal ages: the Mixed group is the youngest (M = 24.7 against 29.0 and 28.4) and includes none of the six listeners over 50.

2.3. Material

A male actor from Zurich in his early twenties read the passage three times [10]. He wrote it himself; it was later adopted as the read text used in the *Dialektatlas* project [11], since read by 1016 Swiss speakers across 125 localities. The six manipulated variables are those described above: /s/-voicing, where Swiss speakers are uniform; <ä>, the stereotype; <ch>, /k/-affrication and /r/, where speakers differ; and <chs> in *sechs* [6], [7]. Swiss-near was intended as the Swiss realization throughout, German-near as the German one, and Mixed as an intermediate keeping the Swiss /r/ quality and /s/. Speaker, text and recording conditions were constant.

2.4. Stimulus verification

Because an actor produced the guises, we verified the manipulation. Target segments were located by forced alignment [12] and coded auditorily, token by token, by the first author, who was not blind to guise. The six manipulated variables give seven coding categories, /r/ being coded twice, for quality and for vocalization: /r/ quality, /r/ vocalization, <ch>, /k/-affrication, /s/-voicing, <ä>-openness and <chs> in *sechs*. Two controls, glottal stop before vowel-initial words and /p, t/ aspiration, both reported as Swiss German differences [4], [5], check whether the performance spills beyond its targets. That makes nine variables and 70 tokens per guise, 210 in all. On every variable the German-variant share is lowest in Swiss-near and highest in German-near (Figure 1), rising from 13% through 59% to 89% overall. The variables <ch>, /k/ and <chs> are categorical at 0 and 100%. Mixed is split on /r/: it keeps the apical quality where the r is consonantal (9 of 10) but vocalizes it postvocalically (13 of 15), which was not intended. The /k/ release duration corroborates the affrication coding, longest in Swiss-near in 4 of 4 tokens (guise means 127, 92 and 80 ms).

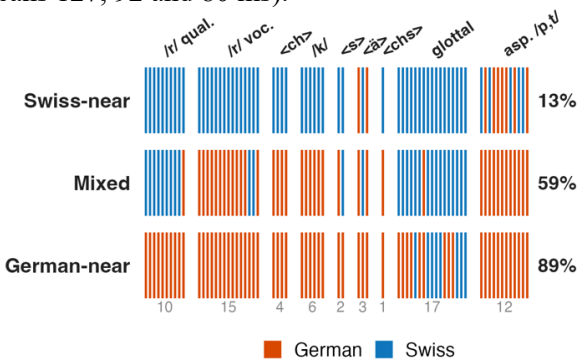

Figure 1. Each cell is one token coded auditorily as the Swiss (blue) or the German (red) variant, grouped by variable, with token counts below each block and the German-variant share at the right.

Four more things did not come out as intended. Swiss-near aspirates /p, t/ in 7 of 12 tokens and uses the norm vowel in 2 of 3 <ä> tokens, expected of a Zurich speaker since the over-open <ä> is now largely a western feature [4]; German-near has a glottal stop before only 9 of 17 vowel-initial words, the one variable on which it falls short of the norm; and Mixed voices /s/ in *sich* though it was intended voiceless.

2.5. Prosodic properties

The guises also differ prosodically; values run Swiss-near, Mixed, German-near: articulation rate is 5.60, 5.85 and 5.34 syll/s, and pauses fill 20.6, 19.6 and 20.4% of each recording, so the difference

Under review. Submitted to ICPhS 2027. Please do not cite without permission.

lies in articulation rather than pausing. Vocalic durational variability is greatest in German-near (nPVI-V 55.6, 52.5, 60.0 [13]; VarcoV 53.8, 52.4, 58.2 [14]) and consonantal variability flat, the direction expected if the German-sounding guise is more stress-timed and the Swiss more syllable-timed [4]. Pitch level is near-identical across the three: the medians lie within 4 Hz and the 10th-to-90th-percentile ranges within a semitone.

2.6. Analysis

Each listener heard one guise at random, so nobody could compare. Ratings are ordered categories, not measurements, so we used cumulative link models for analysis [15]. For each scale we tested whether guise improved the fit (likelihood ratio, with age, gender and education held constant, Holm-corrected across the three; n = 96, two "other" answers on education dropped). Proportional odds held wherever the test could be computed, failed only for age. We then tested whether guise matters more on some scales, fitting all 288 ratings in a cumulative link mixed model [15] with one random intercept per listener, with and without a scale × guise interaction. A linear mixed model [16] splits the variance between guise, listener and background, and returns the same interaction. The interaction was rerun twice: without the 13 listeners who never varied their rating, and without the two duplicate rows. Moderation was tested by adding guise × age, gender and education, Holm-corrected. Descriptively, we also averaged likability and attractiveness into a solidarity composite (α = .80) and report differences in rating points with 95% confidence intervals.

3. RESULTS

3.1. Ratings by guise

The three guises order identically on every scale, with Mixed highest (Figure 2). Means (1–6) for Mixed, Swiss-near and German-near are: competence 4.74, 4.31, 4.03; likability 4.85, 3.88, 3.09; attractiveness 4.00, 3.06, 2.47. Guise is significant on all three scales, decisively on likability and attractiveness (Table 1).

Table 1: Cumulative link models of the three rating scales. Odds ratios are against Mixed; OR < 1 = rated below Mixed (covariates age, gender, education; N = 96).

Scale	Contrast	OR	p
Competence	Swiss-near	0.55	.21
	German-near	0.29	.010
	guise (LR)	$\chi^2(2) = 6.8$	.033
Likability	Swiss-near	0.23	.003
	German-near	0.05	<.0001
	guise (LR)	$\chi^2(2) = 35.5$	<.0001
Attractive-ness	Swiss-near	0.22	.002
	German-near	0.07	<.0001
	guise (LR)	$\chi^2(2) = 29.2$	<.0001

Swiss-near outranks German-near on likability and attractiveness (p = .002, .014) but not on competence (+0.28 [−0.26, 0.82]). German-near never received a top likability rating.

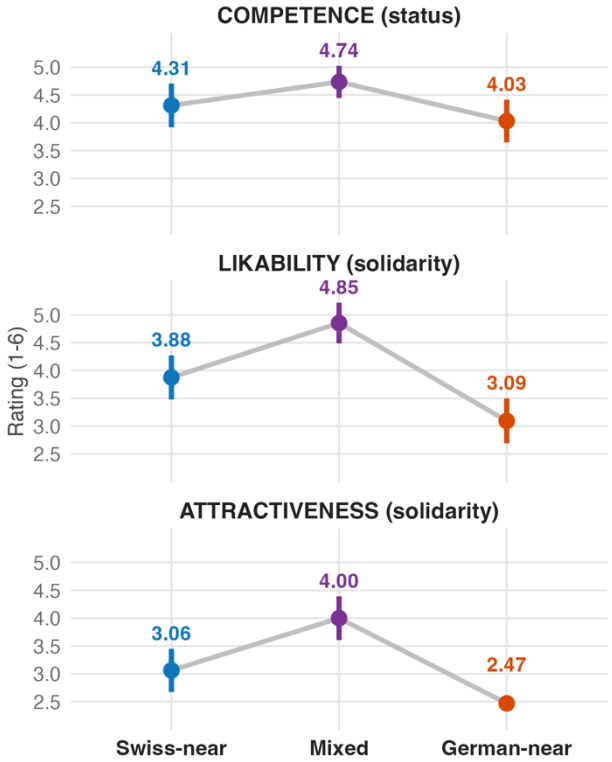

Figure 2. Mean rating (1–6) of one speaker in three guises, on three scales, with 95% confidence intervals. N = 98 listeners, each rating one guise.

3.2. Status and solidarity

Likability and attractiveness correlate strongly (ρ = .69) but only weakly with competence (ρ = .38–.42), so we averaged them into a single solidarity score (Figure 3). On that score Mixed sits 1.65 rating points above German-near (95% CI [1.16, 2.13]) and 0.96 above Swiss-near [0.48, 1.43], and Swiss-near sits 0.69 above German-near [0.20, 1.17]. Guise moves likability and attractiveness alike ($\chi^2(2) = 0.3$, p = .87) and competence

Under review. Submitted to ICPhS 2027. Please do not cite without permission.

differently, which is what licenses pooling the two: across all three scales the interaction is $\chi^2(4) = 13.1$, $p = .011$, and survives dropping either the 13 listeners who gave the same rating on all three scales ($p = .021$) or the two duplicate rows ($p = .019$).

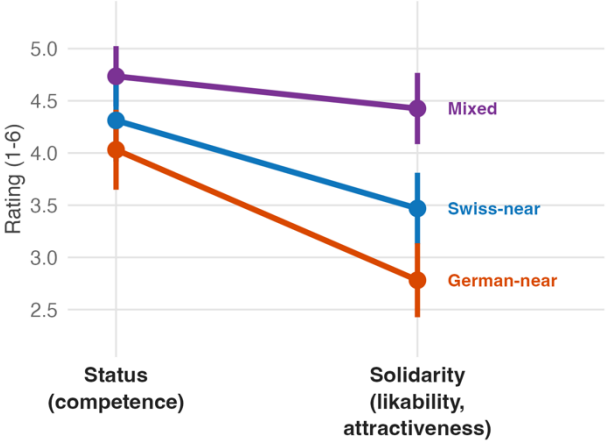

Figure 3. Status and solidarity by guise, with 95% CIs. Solidarity is the mean of likability and attractiveness; the tests are ordinal (Table 1). Each guise was rated by a different group of listeners. N = 98.

3.3. Listener background

Listener background made no difference: age, gender and education neither changed the size of the guise effect (all Holm-corrected $p > .58$) nor predicted how high a listener rated overall, together adding 2.3% of explained variance. Age as a category and region could not be tested: no older listener heard Mixed, and only two listeners came from eastern Switzerland. The highest-rated guise thus also had the youngest listeners, but age did not moderate the guise effect (corrected $p = 1.00$).

4. DISCUSSION

All three hypotheses are supported. The actor's manipulation came out as intended: auditory coding places the readings at 13%, 59% and 89% German variants, a graded contrast. Preference peaks in the middle, supporting H1. Mixed was rated highest on all three scales, and above both Swiss-near and German-near on likability and attractiveness; Swiss-near was in turn rated above German-near. This is what Hove's *Sprachkonvention* predicts: if departure is sanctioned as conceited in one direction and uneducated in the other [5], the best-rated point must lie between them. The bind comes from what Scharloth calls the *asymmetric pluricentricity* of German [17]. The three national standards count as equals, but Germany is the "dominant" center [4]. Approaching the dominant norm promises competence at the cost of loyalty, and refusing it promises belonging at the cost of education; the optimum buys both. That is the ambivalence long reported for Swiss speakers [18], [19].

Guise moved likability and attractiveness far more than competence, supporting H2 and reproducing the status-solidarity split reported for language attitudes [9]. Competence moved too, but less. The stimuli give it a narrower range to work with: all three readings are fluent native standard German, so none is "deficient" in the way, for example, a foreign accent would be. What the variants carry is affiliation, Swiss or German, and that is where the effect is large.

H3 holds in a bounded form: German-near was rated lowest on competence, reliably below Mixed, and we can rule out any advantage over Swiss-near larger than 0.26 points on the six-point scale, so it gained no status. Hove found the same in explicit judgments: German standard was called the better variety, yet two thirds found a Swiss speaking like a German *unsympathisch* [5]. The prestige of the German norm does not transfer to a Swiss speaker adopting it, at least not before Swiss listeners, nor in a reading with no situation attached.

Who the listener was made almost no difference: age, gender and education neither changed the size of the guise effect nor predicted how high anyone rated, though with about 32 listeners per guise a moderate effect would have escaped us. Production is different: among the 1,000 speakers who read this same passage for the *Dialektatlas*, /k/ and <ch> are stratified by region, education, gender, attitudes and political leaning [11]. The social facts that predict how a Swiss person speaks did not predict how they judged. This is consistent with Hove's account: the convention on which variants are acceptable holds for all of German-speaking Switzerland [5], so no group has a separate standard. Listeners do still differ widely, and almost none of that difference is explained by who they are. Since *Hochdeutsch-Ideale* are activated by situation [2] and this recording gave none, each listener had to supply a situation of their own; what divides them may be which one.

Two things limit what the ratings can show. First, German-near struck us as more formal than intended, so we measured F1 to F4: all four sit lower than in the other two guises, consistent with a lowered larynx. The guises differ prosodically too, so listeners may have responded to a persona as much as to the variants [20]. Second, one actor produced all three recordings, and while he is from Zurich, three quarters of the listeners are from Bern, so the estimates hold for this voice and this sample rather than for the varieties.

Under review. Submitted to ICPhS 2027. Please do not cite without permission.

5. REFERENCES

[1] W. Haas, "Die Sprachsituation in der deutschen Schweiz und das Konzept der Diglossie," in Dialekt, Regiolekt und Standardsprache im sozialen und zeitlichen Raum, H. Christen, Ed. Praesens, 2004, pp. 81–110.

[2] M. Guntern, "Dialekt und gesprochene Standardsprache: Wie Laien gesprochenes Schweizerhochdeutsch beurteilen," Sociolinguistica, vol. 26, no. 1, pp. 102–119, 2012.

[3] M. Guntern, "Wie schweizerisch darf des Deutschschweizers Hochdeutsch klingen? Der 'Tatort'-Fall Sabine Timoteo," Schweizerdeutsch, vol. 17, no. 2, pp. 3–6, 2009.

[4] H. Christen and R. Schmidlin, "Deutsch," in Sprachenräume der Schweiz, Band 1: Sprachen, E. Glaser, J. Kabatek, and B. Sonnenhauser, Eds. Narr Francke Attempto, 2024, pp. 21–67, doi: 10.24053/9783381104024.

[5] I. Hove, Die Aussprache der Standardsprache in der deutschen Schweiz, Phonai 47. Niemeyer, 2002.

[6] M. Guntern, "Variables Schweizerhochdeutsch: Eine empirische Untersuchung zur Frage, warum die Standardsprache in der Deutschschweiz wie gesprochen wird und wie gesprochen werden könnte," Ph.D. dissertation, Univ. of Fribourg, Switzerland, 2023, doi: 10.51363/unifr.lth.2023.024.

[7] H. Christen, M. Guntern, I. Hove, and M. Petkova, Hochdeutsch in aller Munde: Eine empirische Untersuchung zur gesprochenen Standardsprache in der Deutschschweiz, ZDL Beiheft 140. Steiner, 2010.

[8] M. Guntern, "Erkennen von Dialekten anhand von gesprochenem Schweizerhochdeutsch," Zeitschrift für Dialektologie und Linguistik, vol. 78, no. 2, pp. 155–187, 2011.

[9] E. B. Ryan and H. Giles, Eds., Attitudes towards Language Variation: Social and Applied Contexts. Edward Arnold, 1982.

[10] C. C. Lanthemann, "Zwischen urbaner Schickeria und Schweizer Eigenart: Der Einfluss des soziodemografischen Hintergrundes auf die Aussprache von Standarddeutsch in der Deutschschweiz," M.A. thesis, Univ. of Bern, Switzerland, 2021.

[11] C. Steiner, D. Studer-Joho, C. Lanthemann, A. Büchler, and A. Leemann, "Sociodemographic and attitudinal effects on dialect speakers' articulation of the standard language: Evidence from German-speaking Switzerland," in Proc. Interspeech 2023, 2023, pp. 3542–3546.

[12] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, "Montreal Forced Aligner: Trainable text-speech alignment using Kaldi," in Proc. Interspeech 2017, 2017, pp. 498–502.

[13] E. Grabe and E. L. Low, "Durational variability in speech and the rhythm class hypothesis," in Laboratory Phonology 7, C. Gussenhoven and N. Warner, Eds. Mouton de Gruyter, 2002, pp. 515–546.

[14] V. Dellwo, "Rhythm and speech rate: A variation coefficient for deltaC," in Language and Language Processing, P. Karnowski and I. Szigeti, Eds. Peter Lang, 2006, pp. 231–241.

[15] R. H. B. Christensen, "ordinal: Regression models for ordinal data," R package version 2026.7-26, 2026. Available: https://CRAN.R-project.org/package=ordinal

[16] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," J. Stat. Softw., vol. 67, no. 1, pp. 1–48, 2015.

[17] J. Scharloth, "Asymmetrische Plurizentrizität und Sprachbewusstsein: Einstellungen der Deutschschweizer zum Standarddeutschen," Zeitschrift für germanistische Linguistik, vol. 33, no. 2–3, pp. 236–267, 2005.

[18] J. Scharloth, "Schweizer Hochdeutsch – schlechtes Hochdeutsch?" in Schweizer Standarddeutsch: Beiträge zur Varietätenlinguistik, C. Dürscheid and M. Businger, Eds. Narr, 2006, pp. 81–96.

[19] S. Hägi and J. Scharloth, "Ist Standarddeutsch für Deutschschweizer eine Fremdsprache? Untersuchungen zu einem Topos des sprachreflexiven Diskurses," Linguistik online, vol. 24, no. 3, pp. 19–47, 2005.

[20] T. Pistor and A. Leemann, "Echoes of implicit bias: Exploring aesthetics and social meanings of Swiss German dialect features," in Proc. Interspeech 2024, 2024, doi: 10.21437/Interspeech.2024-2013.

Under review. Submitted to ICPhS 2027. Please do not cite without permission.