Length & gradience in Dolgan rounding harmony

Deepthi Gopal[1], *Stephen Nichols*[2],
László Károly[1] *& Pavel Iosad*[2]

[1]Uppsala University

[2]University of Edinburgh

4 February 2025
OCP Workshop on Vowel Harmony, Leiden, Netherlands

The diachronic trajectory of vowel harmony

Stability or breakdown?

- ‘Vowel harmony is apparently diachronically-robust.

The diachronic trajectory of vowel harmony

Stability or breakdown?

- **'Vowel harmony is apparently diachronically-robust.**
  - Well-attested on the order of millennia in e.g. Turkic & Uralic (Binnick 1991; Harrison, Dras & Kapicioglu 2006).

The diachronic trajectory of vowel harmony

Stability or breakdown?

- **'Vowel harmony is apparently diachronically-robust.**
  - Well-attested on the order of millennia in e.g. Turkic & Uralic (Binnick 1991; Harrison, Dras & Kapicioglu 2006).
  - Binnick (1991): vowel harmony is unaffected by borrowing. (!)

The diachronic trajectory of vowel harmony

Stability or breakdown?

- **'Vowel harmony is apparently diachronically-robust.**
  - Well-attested on the order of millennia in e.g. Turkic & Uralic (Binnick 1991; Harrison, Dras & Kapicioglu 2006).
  - Binnick (1991): vowel harmony is unaffected by borrowing. (!)
- **At the same time, we know that harmony decay exists!**

The diachronic trajectory of vowel harmony

Stability or breakdown?

- **'Vowel harmony is apparently diachronically-robust.**
  - Well-attested on the order of millennia in e.g. Turkic & Uralic (Binnick 1991; Harrison, Dras & Kapicioglu 2006).
  - Binnick (1991): vowel harmony is unaffected by borrowing. (!)
- **At the same time, we know that harmony decay exists!**
  - Growing literature on harmony loss (Kavitskaya 2013; McCollum 2015; McCollum & Chen 2020; McCollum 2020; Sandstedt 2020; Kavitskaya & McCollum 2023).

The diachronic trajectory of vowel harmony

Stability or breakdown?

- **'Vowel harmony is apparently diachronically-robust.**
  - Well-attested on the order of millennia in e.g. Turkic & Uralic (Binnick 1991; Harrison, Dras & Kapicioglu 2006).
  - Binnick (1991): vowel harmony is unaffected by borrowing. (!)
- **At the same time, we know that harmony decay exists!**
  - Growing literature on harmony loss (Kavitskaya 2013; McCollum 2015; McCollum & Chen 2020; McCollum 2020; Sandstedt 2020; Kavitskaya & McCollum 2023).
  - Binnick (1991): vowel harmony is inherently unstable on its own account. (!)

The diachronic trajectory of vowel harmony

Stability or breakdown?

- **'Vowel harmony is apparently diachronically-robust.**
  - Well-attested on the order of millennia in e.g. Turkic & Uralic (Binnick 1991; Harrison, Dras & Kapicioglu 2006).
  - Binnick (1991): vowel harmony is unaffected by borrowing. (!)
- **At the same time, we know that harmony decay exists!**
  - Growing literature on harmony loss (Kavitskaya 2013; McCollum 2015; McCollum & Chen 2020; McCollum 2020; Sandstedt 2020; Kavitskaya & McCollum 2023).
  - Binnick (1991): vowel harmony is inherently unstable on its own account. (!)
- **What drives harmony decay?**

The diachronic trajectory of vowel harmony

Stability or breakdown?

- **'Vowel harmony is apparently diachronically-robust.**
  - Well-attested on the order of millennia in e.g. Turkic & Uralic (Binnick 1991; Harrison, Dras & Kapicioglu 2006).
  - Binnick (1991): vowel harmony is unaffected by borrowing. (!)
- **At the same time, we know that harmony decay exists!**
  - Growing literature on harmony loss (Kavitskaya 2013; McCollum 2015; McCollum & Chen 2020; McCollum 2020; Sandstedt 2020; Kavitskaya & McCollum 2023).
  - Binnick (1991): vowel harmony is inherently unstable on its own account. (!)
- **What drives harmony decay?**
  - Two main sources of harmony loss recur in the literature.

The diachronic trajectory of vowel harmony

Stability or breakdown?

- **'Vowel harmony is apparently diachronically-robust.**
  - Well-attested on the order of millennia in e.g. Turkic & Uralic (Binnick 1991; Harrison, Dras & Kapicioglu 2006).
  - Binnick (1991): vowel harmony is unaffected by borrowing. (!)
- **At the same time, we know that harmony decay exists!**
  - Growing literature on harmony loss (Kavitskaya 2013; McCollum 2015; McCollum & Chen 2020; McCollum 2020; Sandstedt 2020; Kavitskaya & McCollum 2023).
  - Binnick (1991): vowel harmony is inherently unstable on its own account. (!)
- **What drives harmony decay?**
  - Two main sources of harmony loss recur in the literature.
    - Phonologisation of vowel reduction.
    - 'Language contact'.

Harmony decay & harmony retention

Why decay?

- **Vowel reduction.**

Harmony decay & harmony retention

Why decay?

- **Vowel reduction.**
  - On one hand, the predictability induced by harmony may permit reduction. (Binnick 1991; McCollum 2020)

Harmony decay & harmony retention

Why decay?

- **Vowel reduction.**
  - On one hand, the predictability induced by harmony may permit reduction. (Binnick 1991; McCollum 2020)
  - On the other hand, sufficient vowel reduction may obscure the phonetic effects of harmony, especially **rounding** harmony, due to crowding in the middle of the acoustic space. (Binnick 1991)

Harmony decay & harmony retention

Why decay?

- **Vowel reduction.**
  - On one hand, the predictability induced by harmony may permit reduction. (Binnick 1991; McCollum 2020)
  - On the other hand, sufficient vowel reduction may obscure the phonetic effects of harmony, especially **rounding** harmony, due to crowding in the middle of the acoustic space. (Binnick 1991)
    - Reduction claimed to partially contribute to decay in e.g. Nivkh (Shiraishi & Botma 2017), Sanjiazi Manchu (Li 1996).

Harmony decay & harmony retention

Why decay?

- **Vowel reduction.**
  - On one hand, the predictability induced by harmony may permit reduction. (Binnick 1991; McCollum 2020)
  - On the other hand, sufficient vowel reduction may obscure the phonetic effects of harmony, especially **rounding** harmony, due to crowding in the middle of the acoustic space. (Binnick 1991)
    - Reduction claimed to partially contribute to decay in e.g. Nivkh (Shiraishi & Botma 2017), Sanjiazi Manchu (Li 1996).
  - Pearce (2008, 2012) claims that, in Kera (Chadic), harmonic vowels resist reduction in quality but *non*-harmonic vowels do not (contra McCollum 2020!).

Harmony decay & harmony retention

Why decay?

- **Vowel reduction.**
  - On one hand, the predictability induced by harmony may permit reduction. (Binnick 1991; McCollum 2020)
  - On the other hand, sufficient vowel reduction may obscure the phonetic effects of harmony, especially **rounding** harmony, due to crowding in the middle of the acoustic space. (Binnick 1991)
    - Reduction claimed to partially contribute to decay in e.g. Nivkh (Shiraishi & Botma 2017), Sanjiazi Manchu (Li 1996).
  - Pearce (2008, 2012) claims that, in Kera (Chadic), harmonic vowels resist reduction in quality but *non*-harmonic vowels do not (contra McCollum 2020!).
  - Reduction also implicated in the emergence of harmony (e.g. Hyman 2002).

Harmony decay & harmony retention

Why decay?

- **Vowel reduction.**
  - **Centralisation** of the vowel space reported for several Turkic languages: Kyrgyz (McCollum 2020), Tatar (Conklin & Dmitrieva 2018), Uyghur (McCollum, Durvasula & Abudushalamu 2024), with or without attendant harmony decay.

Harmony decay & harmony retention

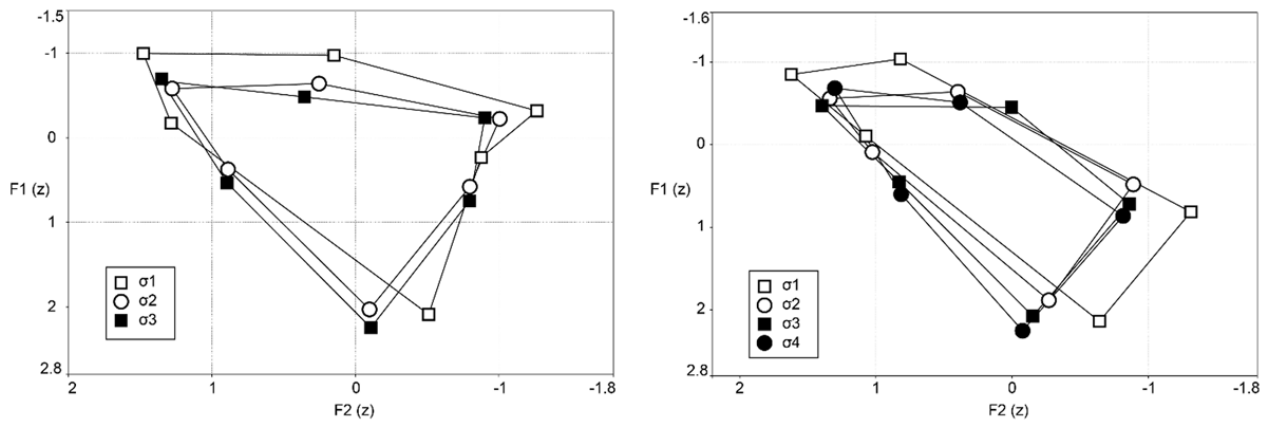
Why decay?

- **Vowel reduction.**
  - **Centralisation** of the vowel space reported for several Turkic languages: Kyrgyz (McCollum 2020), Tatar (Conklin & Dmitrieva 2018), Uyghur (McCollum, Durvasula & Abudushalamu 2024), with or without attendant harmony decay.
  - For Sakha, the closest relative of Dolgan, Chan & Kuang (2023) report that 'target [short] vowels are gradually centralized when further away from the trigger'.

Harmony decay & harmony retention

Why decay?

- **Vowel reduction.**
  - **Centralisation** of the vowel space reported for several Turkic languages: Kyrgyz (McCollum 2020), Tatar (Conklin & Dmitrieva 2018), Uyghur (McCollum, Durvasula & Abudushalamu 2024), with or without attendant harmony decay.
  - For Sakha, the closest relative of Dolgan, Chan & Kuang (2023) report that 'target [short] vowels are gradually centralized when further away from the trigger'.



McCollum (2020): Vowel reduction in Kyrgyz.

Harmony decay & harmony retention

Why decay?

- **Language contact.**
    - Both decay and emergence have been attributed to contact.

Harmony decay & harmony retention

Why decay?

- **Language contact.**
  - Both decay and emergence have been attributed to contact.
  - Emergence. Kavitskaya & McCollum (2023): spread of RH to non-high vowels in Siberian & Central Asian Turkic is due to Mongolic (in which only non-high vowels undergo RH).

Harmony decay & harmony retention

Why decay?

- **Language contact.**
  - Both decay and emergence have been attributed to contact.
  - Emergence. Kavitskaya & McCollum (2023): spread of RH to non-high vowels in Siberian & Central Asian Turkic is due to Mongolic (in which only non-high vowels undergo RH).
  - Emergence Dawkins & Halliday (1916) in McCollum (2024): VH innovated in Greek varieties of Turkey due to Turkish.

Harmony decay & harmony retention

Why decay?

- **Language contact.**
  - Both decay and emergence have been attributed to contact.
  - Emergence. Kavitskaya & McCollum (2023): spread of RH to non-high vowels in Siberian & Central Asian Turkic is due to Mongolic (in which only non-high vowels undergo RH).
  - Emergence Dawkins & Halliday (1916) in McCollum (2024): VH innovated in Greek varieties of Turkey due to Turkish.
  - Decay. Sjoberg (1963); Harrison, Dras & Kapicioglu (2006): Uzbek lost harmony due to long-term contact with Persian.

Harmony decay & harmony retention

Why decay?

- **Language contact.**
  - Both decay and emergence have been attributed to contact.
  - Emergence. Kavitskaya & McCollum (2023): spread of RH to non-high vowels in Siberian & Central Asian Turkic is due to Mongolic (in which only non-high vowels undergo RH).
  - Emergence Dawkins & Halliday (1916) in McCollum (2024): VH innovated in Greek varieties of Turkey due to Turkish.
  - Decay. Sjoberg (1963); Harrison, Dras & Kapicioglu (2006): Uzbek lost harmony due to long-term contact with Persian.
  - Decay. Bobaljik (2018): harmony decay in Itelmen (Chukotko-Kamchatkan) depends crucially on both *structural factors* (vowel merger) and *borrowing* from Russian.
    - Closely-related Chukchi retains harmony for structural reasons.

Harmony decay & harmony retention

Why not just ***not*** *decay?*

- **Language contact.**
  - In fact, many languages with iterative **rounding harmony** (Turkic, Uralic, Tungusic, Mongolic) show:

Harmony decay & harmony retention

*Why not just **not** decay?*

- **Language contact.**
  - In fact, many languages with iterative **rounding harmony** (Turkic, Uralic, Tungusic, Mongolic) show:
    - Contact with one another, often over a prolonged period.

Harmony decay & harmony retention

Why not just ***not*** *decay?*

- **Language contact.**
  - In fact, many languages with iterative **rounding harmony** (Turkic, Uralic, Tungusic, Mongolic) show:
    - Contact with one another, often over a prolonged period.
    - Contact with Russian, a language with no harmony, unpredictable stress, and strong vowel reduction in unstressed syllables.

Harmony decay & harmony retention

*Why not just **not** decay?*

- **Language contact.**
  - In fact, many languages with iterative **rounding harmony** (Turkic, Uralic, Tungusic, Mongolic) show:
    - Contact with one another, often over a prolonged period.
    - Contact with Russian, a language with no harmony, unpredictable stress, and strong vowel reduction in unstressed syllables.
- **Question.** What drives harmony retention in the face of conditions that favour decay?

Dolgan

The picture across Turkic

- Where does Dolgan fit in?

Dolgan

The picture across Turkic

- Where does Dolgan fit in?

Language	Harmony loss?	Centralisation?	Russian contact?	Phonemic length?	Sources
Crimean Tatar	yes (rounding)	yes	yes	no	McCollum & Kavitskaya 2022
Kazakh	yes (rounding)	yes	yes	no	McCollum 2015
Kyrgyz	no	yes	yes	**yes**	McCollum 2020
Sakha	no	yes	yes	**yes**	Chan & Kuang 2023
(Kazan) Tatar	yes (rounding)	yes	yes	no	Conklin & Dmitrieva 2018
Uighur	no	yes	**no**	no	McCollum, Durvasula & Abudushalamu 2024
Uzbek	yes (total)	yes	*yes**	no	Sjoberg 1963; Harrison, Dras & Kapicioglu 2006

* Uzbek VH loss predates Russian contact, but is due to contact with a different non-VH lg.

Dolgan

The language

Dolgan territory (reproduced from Däbritz 2022: 4).

Dolgan

The language

Dolgan settlements (reproduced from Däbritz 2022: 5).

Dolgan

The language

- Endangered Turkic language of the Taimyr peninsula (< 1000 speakers).

Dolgan

The language

- Endangered Turkic language of the Taimyr peninsula (< 1000 speakers).
- Closely related to Sakha (Yakut); historically considered a dialect.

Dolgan

The language

- Endangered Turkic language of the Taimyr peninsula (< 1000 speakers).
- Closely related to Sakha (Yakut); historically considered a dialect.
    - Differs from Sakha both structurally and sociolinguistically; much more minoritised, much less formal education available in Dolgan, less written standard.

Dolgan

The language

- Endangered Turkic language of the Taimyr peninsula (< 1000 speakers).
- Closely related to Sakha (Yakut); historically considered a dialect.
    - Differs from Sakha both structurally and sociolinguistically; much more minoritised, much less formal education available in Dolgan, less written standard.
- Essentially all speakers are now bilingual in Russian.

Dolgan

The language

- Endangered Turkic language of the Taimyr peninsula (< 1000 speakers).
- Closely related to Sakha (Yakut); historically considered a dialect.
    - Differs from Sakha both structurally and sociolinguistically; much more minoritised, much less formal education available in Dolgan, less written standard.
- Essentially all speakers are now bilingual in Russian.
- Long-term contact with Evenki (Tungusic) in particular (also Mongolic, Samoyedic).

Dolgan

The language

- Endangered Turkic language of the Taimyr peninsula (< 1000 speakers).
- Closely related to Sakha (Yakut); historically considered a dialect.
    - Differs from Sakha both structurally and sociolinguistically; much more minoritised, much less formal education available in Dolgan, less written standard.
- Essentially all speakers are now bilingual in Russian.
- Long-term contact with Evenki (Tungusic) in particular (also Mongolic, Samoyedic).
- Next to no phonological literature, but thorough recent descriptive grammar (Däbritz 2022).

Dolgan

The language

- Dolgan has a familiar-looking vowel inventory:

Dolgan

The language

- Dolgan has a familiar-looking vowel inventory:

High	i iː	y yː	ɨ ɨː	u uː
Non-high	(e∼ɛ) eː	(ø∼œ) øː	a aː	(o∼ɔ) oː
Diphthong	i̯e̯	y̯œ̯	ɨ̯a̯	u̯o̯

Dolgan

The language

- Dolgan has a familiar-looking vowel inventory:

High	i iː	y yː	ɨ ɨː	u uː
Non-high	(e∼ɛ) eː	(ø∼œ) øː	a aː	(o∼ɔ) oː
Diphthong	ie̯	yœ̯	ɨa̯	uo̯

- The diphthongs /ie, yœ̯, uo/ historically descend mainly from long *mid* vowels; occasionally also from lenitions in VCV sequences (especially of velars).
  - E.g. **bēš* > /bie̯s/ ‘five’, **tȫrt* > /tyœ̯rt/ ‘four’, **ōn* > /uo̯n/ ‘ten’.
  - Long mid vowels re-innovated since this development.

Dolgan

The language

- Dolgan has a familiar-looking vowel inventory:

High	i iː	y yː	ɨ ɨː	u uː
Non-high	(e∼ɛ) eː	(ø∼œ) øː	a aː	(o∼ɔ) oː
Diphthong	ie̯	yœ̯	ɨa̯	uo̯

- The diphthongs /ie, yœ̯, uo/ historically descend mainly from long *mid* vowels; occasionally also from lenitions in VCV sequences (especially of velars).
  - E.g. **bēš* > /bie̯s/ ‘five’, **tȫrt* > /tyœ̯rt/ ‘four’, **ōn* > /uo̯n/ ‘ten’.
  - Long mid vowels re-innovated since this development.
- /ɨa̯/ seems to derive from /aCI/ sequences.
  - E.g. **tabul* > /tɨa̯l/ ‘wind’, **biagïr* > /bɨa̯r/ ‘liver’.

Dolgan

The language

- As is typical for a Turkic language, Dolgan exhibits both backness and rounding harmonies.

Dolgan

The language

- As is typical for a Turkic language, Dolgan exhibits both backness and rounding harmonies.
    - Backness harmony: *Vowels agree for backness.*

Dolgan

The language

- As is typical for a Turkic language, Dolgan exhibits both backness and rounding harmonies.
  - Backness harmony: *Vowels agree for backness.*
  - Rounding harmony: *High vowels only trigger rounding in other high vowels, non-high vowels trigger rounding in all targets.*

Dolgan

The language

- As is typical for a Turkic language, Dolgan exhibits both backness and rounding harmonies.
  - Backness harmony: *Vowels agree for backness.*
  - Rounding harmony: *High vowels only trigger rounding in other high vowels, non-high vowels trigger rounding in all targets.*

	High trigger	Non-high trigger
High target	/uː-nI/ uːnu ‘water-ACC’ /yŋyː-nI/ yŋyːny ‘spear-ACC’	/ogo-nI/ ogonu ‘children-ACC’ /børø-nI/ børøny ‘wolf-ACC’
Non-high target	/uː-lAr/ uːlar ‘water-PL’ /yŋyː-lAr/ yŋyːler ‘spear-PL’	/ok-lAr/ oktor ‘arrow-PL’ /børø-lAr/ børølør ‘wolf-PL’

Dolgan

The language

- As is typical for a Turkic language, Dolgan exhibits both backness and rounding harmonies.
  - Backness harmony: *Vowels agree for backness.*
  - Rounding harmony: *High vowels only trigger rounding in other high vowels, non-high vowels trigger rounding in all targets.*

	High trigger	Non-high trigger
High target	/uː-nI/ uːnu ‘water-ACC’ /yŋyː-nI/ yŋyːny ‘spear-ACC’	/ogo-nI/ ogonu ‘children-ACC’ /børø-nI/ børøny ‘wolf-ACC’
Non-high target	/uː-lAr/ uːlar ‘water-PL’ /yŋyː-lAr/ yŋyːler ‘spear-PL’	/ok-lAr/ oktor ‘arrow-PL’ /børø-lAr/ børølør ‘wolf-PL’

- Diphthongs behave like high vowels (Däbritz 2022: 55), as in Sakha (Chan & Kuang 2023: 3296).

Dolgan

The corpus

- **INEL corpus of Dolgan.** (Däbritz, Kudryakova & Stapert 2022; Däbritz 2020): audio, time-aligned transcription, glossing etc.

Dolgan

The corpus

- **INEL corpus of Dolgan.** (Däbritz, Kudryakova & Stapert 2022; Däbritz 2020): audio, time-aligned transcription, glossing etc.
  - We use most, but not quite all of the corpus (technical difficulties parsing; some missing audio).

Dolgan

The corpus

- **INEL corpus of Dolgan.** (Däbritz, Kudryakova & Stapert 2022; Däbritz 2020): audio, time-aligned transcription, glossing etc.
  - We use most, but not quite all of the corpus (technical difficulties parsing; some missing audio).
  - 42 speakers (30 F, 12 M), birthyears 1912–2009 (median 1936).

Dolgan

The corpus

- **INEL corpus of Dolgan.** (Däbritz, Kudryakova & Stapert 2022; Däbritz 2020): audio, time-aligned transcription, glossing etc.
  - We use most, but not quite all of the corpus (technical difficulties parsing; some missing audio).
  - 42 speakers (30 F, 12 M), birthyears 1912–2009 (median 1936).
  - C. 11 hours 41 min of data, median talk time 5 min, range 2–60 min (SD 12).

Dolgan

The corpus

- **INEL corpus of Dolgan.** (Däbritz, Kudryakova & Stapert 2022; Däbritz 2020): audio, time-aligned transcription, glossing etc.
  - We use most, but not quite all of the corpus (technical difficulties parsing; some missing audio).
  - 42 speakers (30 F, 12 M), birthyears 1912–2009 (median 1936).
  - C. 11 hours 41 min of data, median talk time 5 min, range 2–60 min (SD 12).
  - 57,831 spoken words, 15,877 types, 6,466 unique stems.

Dolgan

The corpus

- **INEL corpus of Dolgan.** (Däbritz, Kudryakova & Stapert 2022; Däbritz 2020): audio, time-aligned transcription, glossing etc.
  - We use most, but not quite all of the corpus (technical difficulties parsing; some missing audio).
  - 42 speakers (30 F, 12 M), birthyears 1912–2009 (median 1936).
  - C. 11 hours 41 min of data, median talk time 5 min, range 2–60 min (SD 12).
  - 57,831 spoken words, 15,877 types, 6,466 unique stems.
- Montreal Forced Aligner (McAuliffe et al. 2017) for segmentation; trained a customised model for Dolgan; manual checking in progress.

Dolgan

The corpus

- **Audio data.**

Dolgan

The corpus

- **Audio data.**
  - 148,092 vowels (137,576 monophthongs, 10,516 diphthongs).

Dolgan

The corpus

- **Audio data.**
  - 148,092 vowels (137,576 monophthongs, 10,516 diphthongs).
  - F1, F2, F3 taken (automatically) at 25%, 50%, 75% of vowel duration. Midpoint or average shown for monophthongs, trajectory shown for diphthongs.

Dolgan

The corpus

- **Audio data.**
  - 148,092 vowels (137,576 monophthongs, 10,516 diphthongs).
  - F1, F2, F3 taken (automatically) at 25%, 50%, 75% of vowel duration. Midpoint or average shown for monophthongs, trajectory shown for diphthongs.
  - *z*-score normalised.

Dolgan

The corpus

- **Audio data.**
  - 148,092 vowels (137,576 monophthongs, 10,516 diphthongs).
  - F1, F2, F3 taken (automatically) at 25%, 50%, 75% of vowel duration. Midpoint or average shown for monophthongs, trajectory shown for diphthongs.
  - *z*-score normalised.
  - Data tagged by us for: details of vowel (phonemic length, rounding, fronting, height); syllable count; root vs. affix status, along with morphological material; preceding & following consonantal context; underspecification.

Dolgan

Language contact

- **Can we evaluate how much borrowing we see?**
  - Focusing on lexical borrowing; plenty of morphological borrowing from Mongolic and Evenki, largely early.

Borrowing source	tokens	% tokens	stems	% stems
Native lexicon	45030	77.9	3274	51.2
Russian	9097	15.7	2765	43.2
Mongolic	3318	5.74	447	3.85
Evenki	359	0.621	171	1.38
Sakha	13	< 0.1	8	0.124
Nganasan	10	< 0.1	5	< 0.1

Dolgan

Language contact

- **Can we evaluate how much borrowing we see?**
  - We can also break this down by genre: `conv` conversational data, `flk` folklore, `nar` narrative, `misc` miscellaneous.
  - For all tokens:

Borrowing source	conv		flk		nar		misc	
	n	%	n	%	n	%	n	%
Native lexicon	12625	74.6	7440	83.9	24887	77.9	78	78.0
Russian	3179	18.8	903	10.2	5009	15.7	6	6.0
Mongolic	1064	6.3	429	4.8	1809	5.7	16	16.0
Evenki	45	0.3	93	1.1	221	0.7	0	0.0
Sakha	8	0.0	0	0.0	5	0.0	0	0.0
Nganasan	3	0.0	0	0.0	7	0.0	0	0.0

Dolgan

Language contact

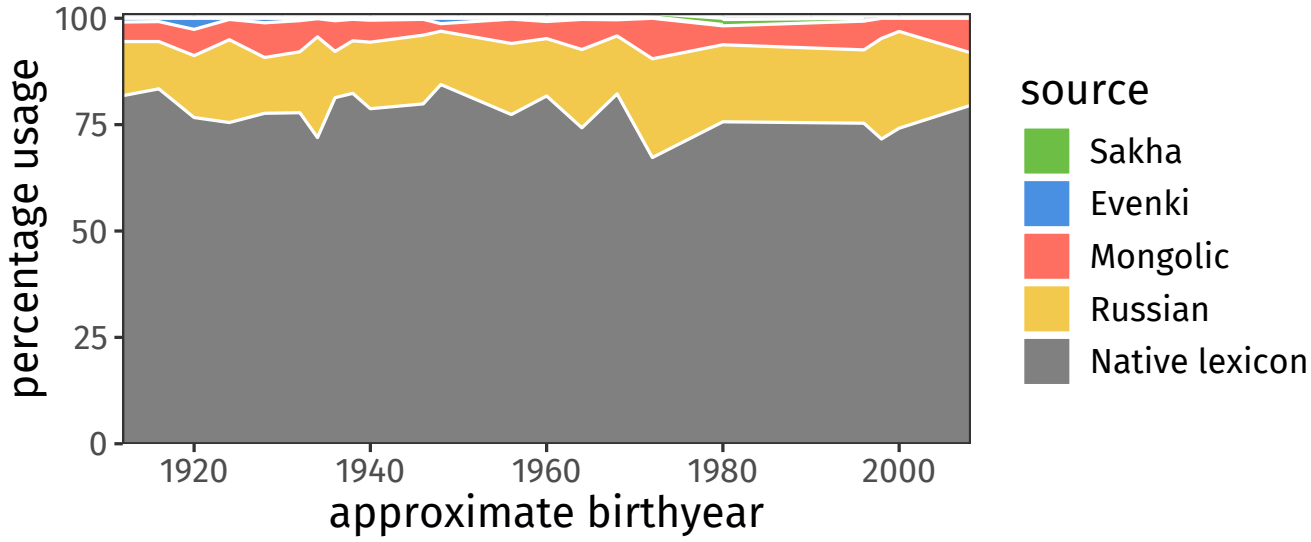
- **Can we evaluate how much borrowing we see?**
  - We can also break this down by genre: `conv` conversational data, `flk` folklore, `nar` narrative, `misc` miscellaneous.
  - And for unique stems:

Borrowing source	conv		flk		nar		misc	
	n	%	n	%	n	%	n	%
Native lexicon	1339	48.7	1143	78.9	2433	54.9	55	80.9
Russian	1270	46.2	184	12.7	1734	39.2	4	5.88
Mongolic	113	4.1	91	6.3	189	4.2	9	13.4
Evenki	21	0.8	30	2.1	61	1.4	0	0.0
Sakha	6	0.2	0	0.0	3	0.1	0	0.0
Nganasan	2	0.1	0	0.0	5	0.1	0	0.0

Dolgan

Language contact

- **Can we evaluate how much borrowing we see?**
  - *Inter-speaker variation?*



Some variation, no statistically-significant trend.

Dolgan

*Language contact & **disharmony***

- **How much disharmony do we see?**
- Overall:

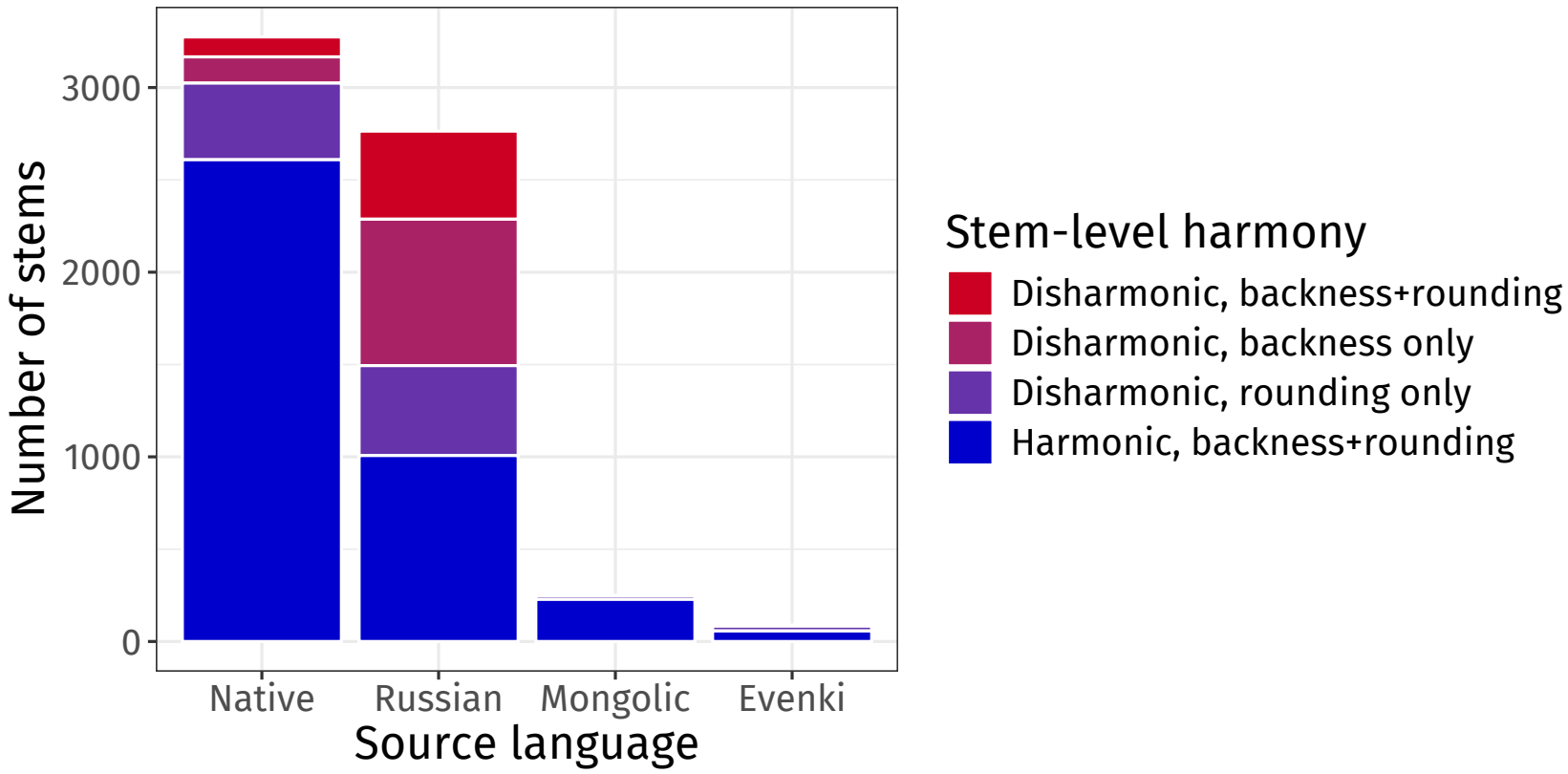
	tokens	% tokens	stems	% stems
Harmonic, backness+rounding	50575	87.5	3804	60.9
Disharmonic, rounding only	4231	7.32	920	14.9
Disharmonic, backness only	1891	3.27	932	14.9
Disharmonic, backness+rounding	1134	1.96	584	9.35

Dolgan

Language contact & disharmony

- How much disharmony do we see?

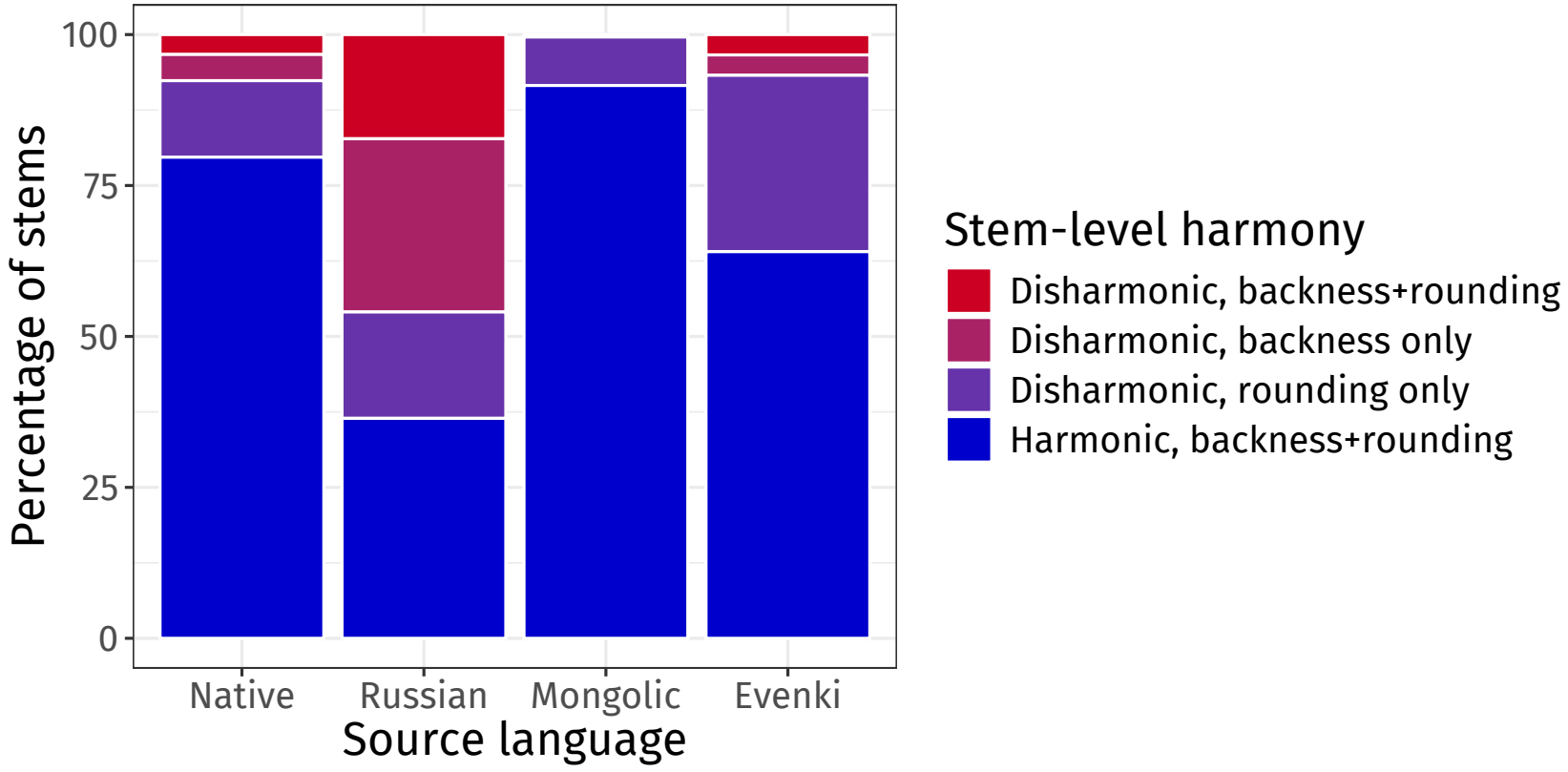
  By source:



Dolgan

Language contact & disharmony

- How much disharmony do we see?

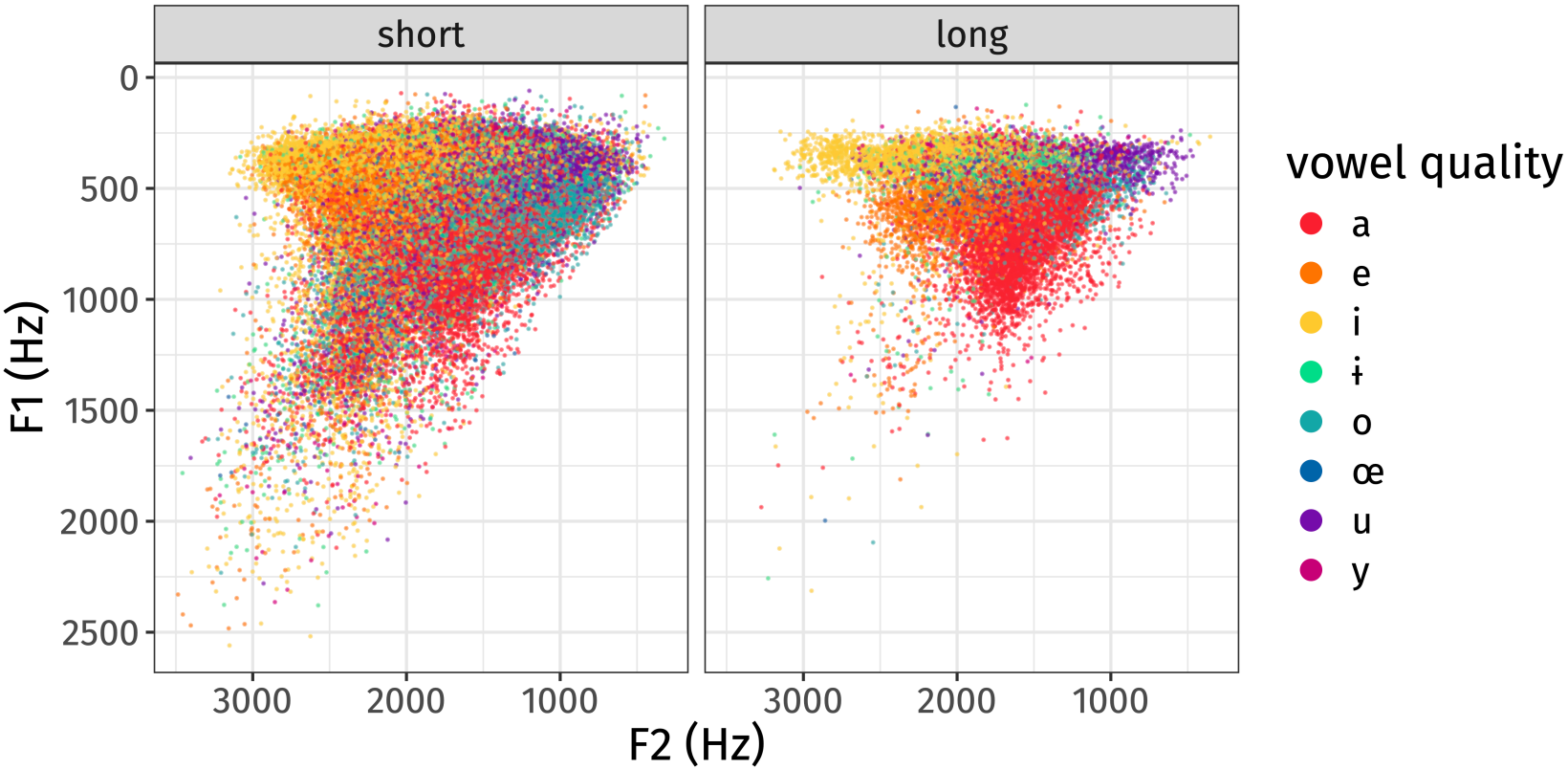


Dolgan

*Language contact & **disharmony***

- **How much disharmony do we see?**
- Plenty of disharmony, mostly driven by Russian borrowing.

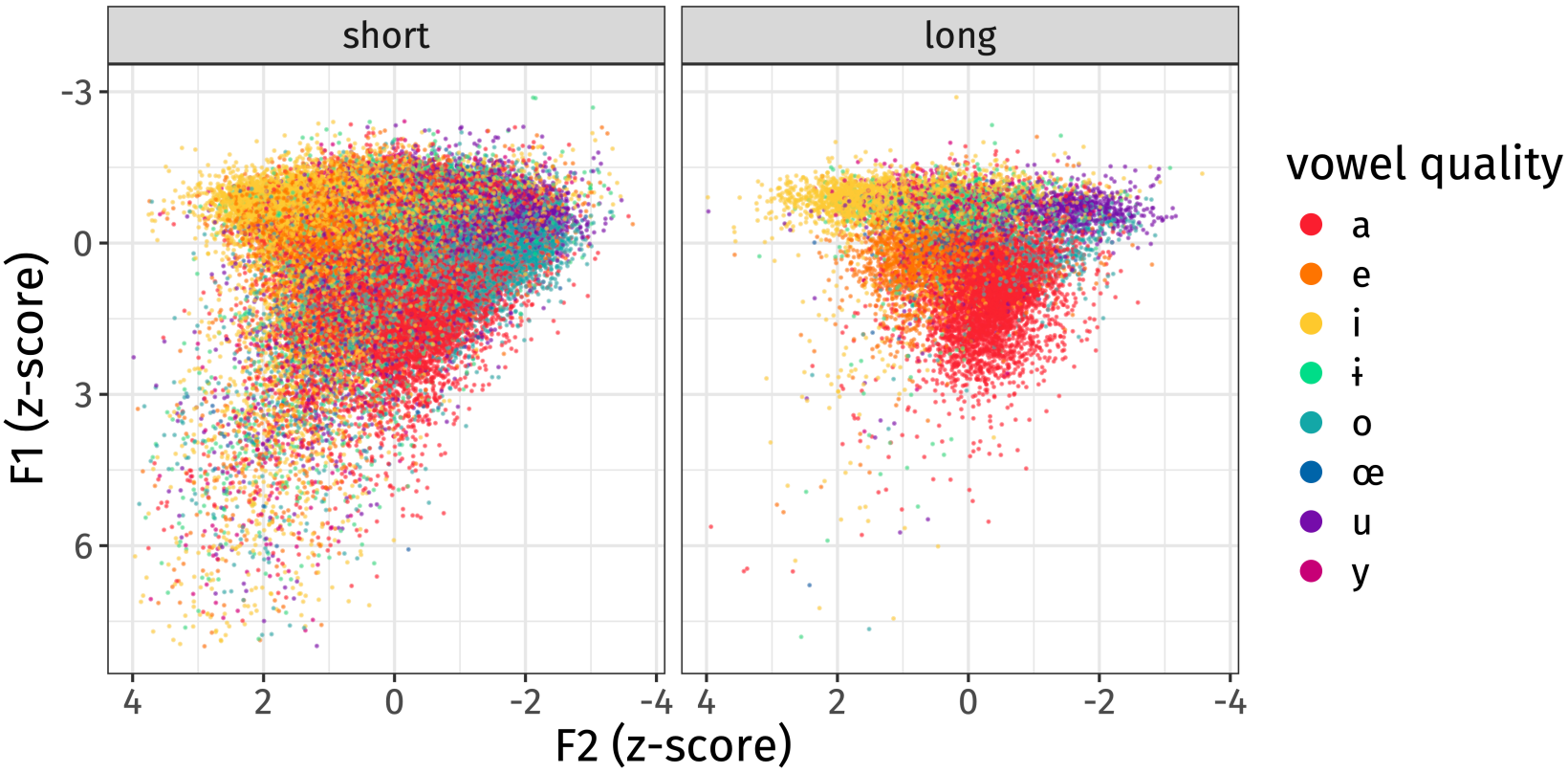
Data quality?



Raw data.

(Predictably,) **long vowels** look reasonably good already; a lot of alignment & measurement error in the **short vowels**.

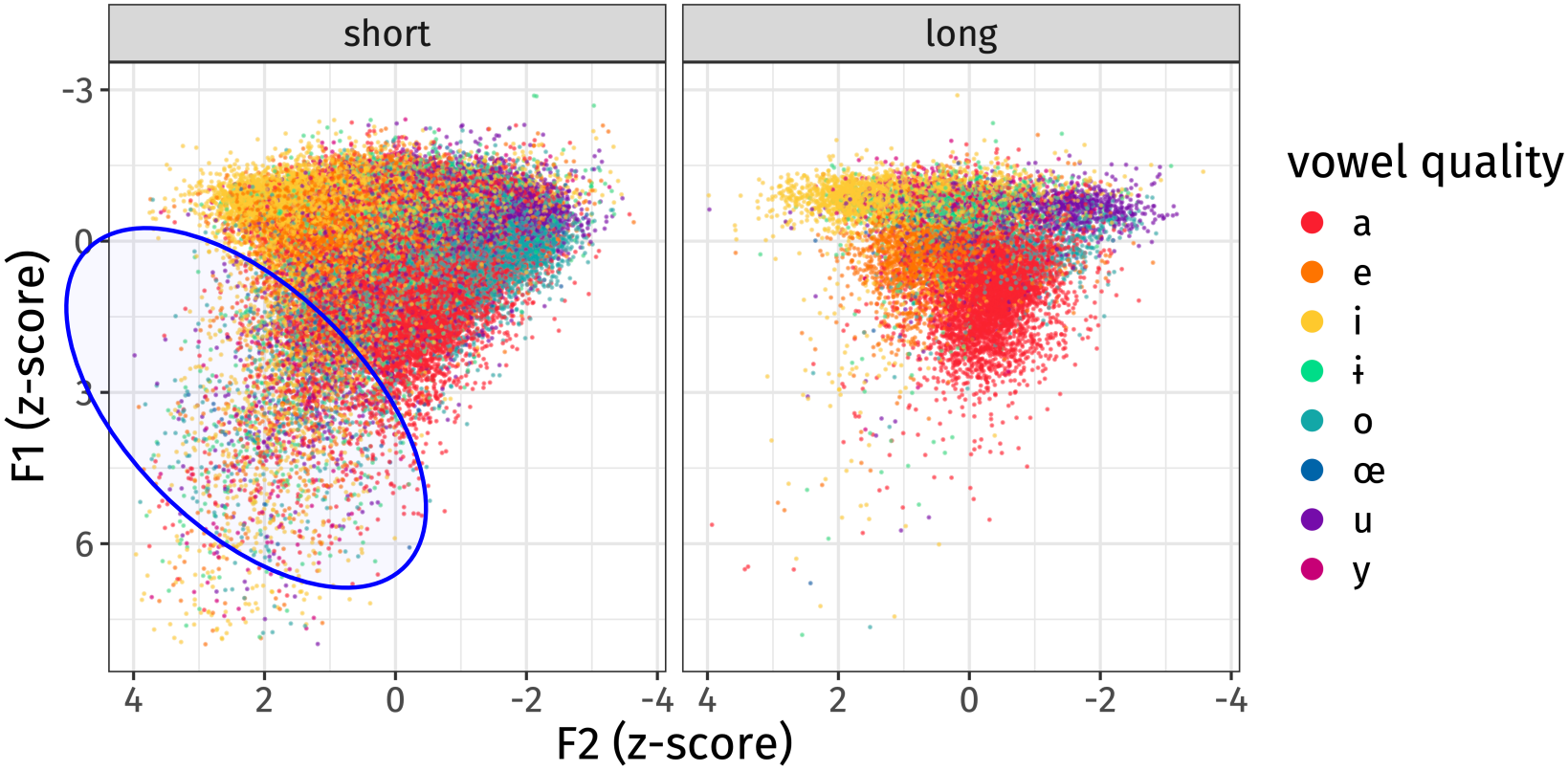
Data quality?



Uncorrected, *z*-score-normalised data.

Normalisation removes some bimodality (due to speaker sex) esp. in the long vowels, but short vowel error remains significant..

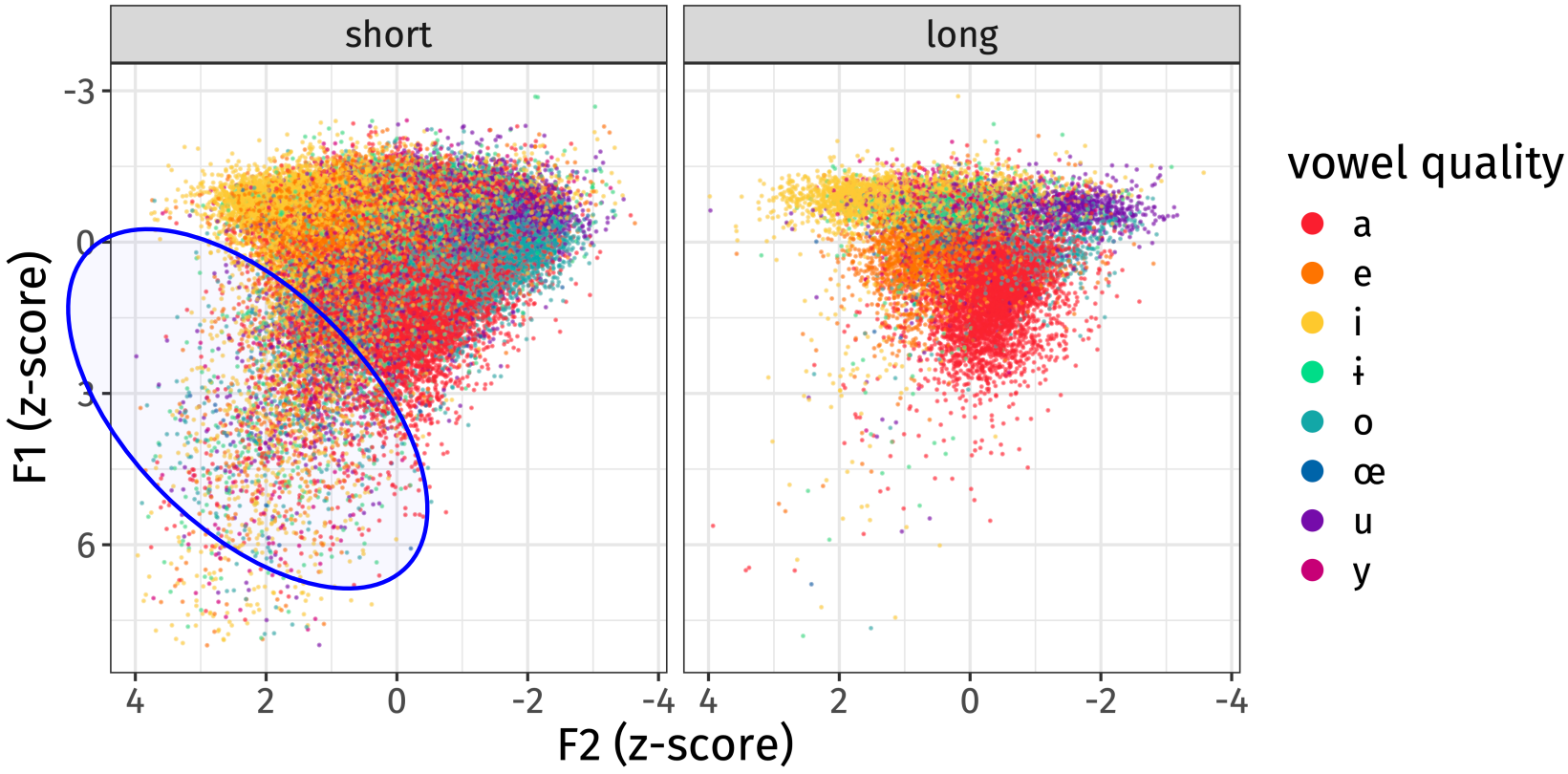
Data quality?



Uncorrected, *z*-score-normalised data.

5-6% (23975/444276 measurements) in 'physically impossible' range, presumably more error within plausible vowel space.

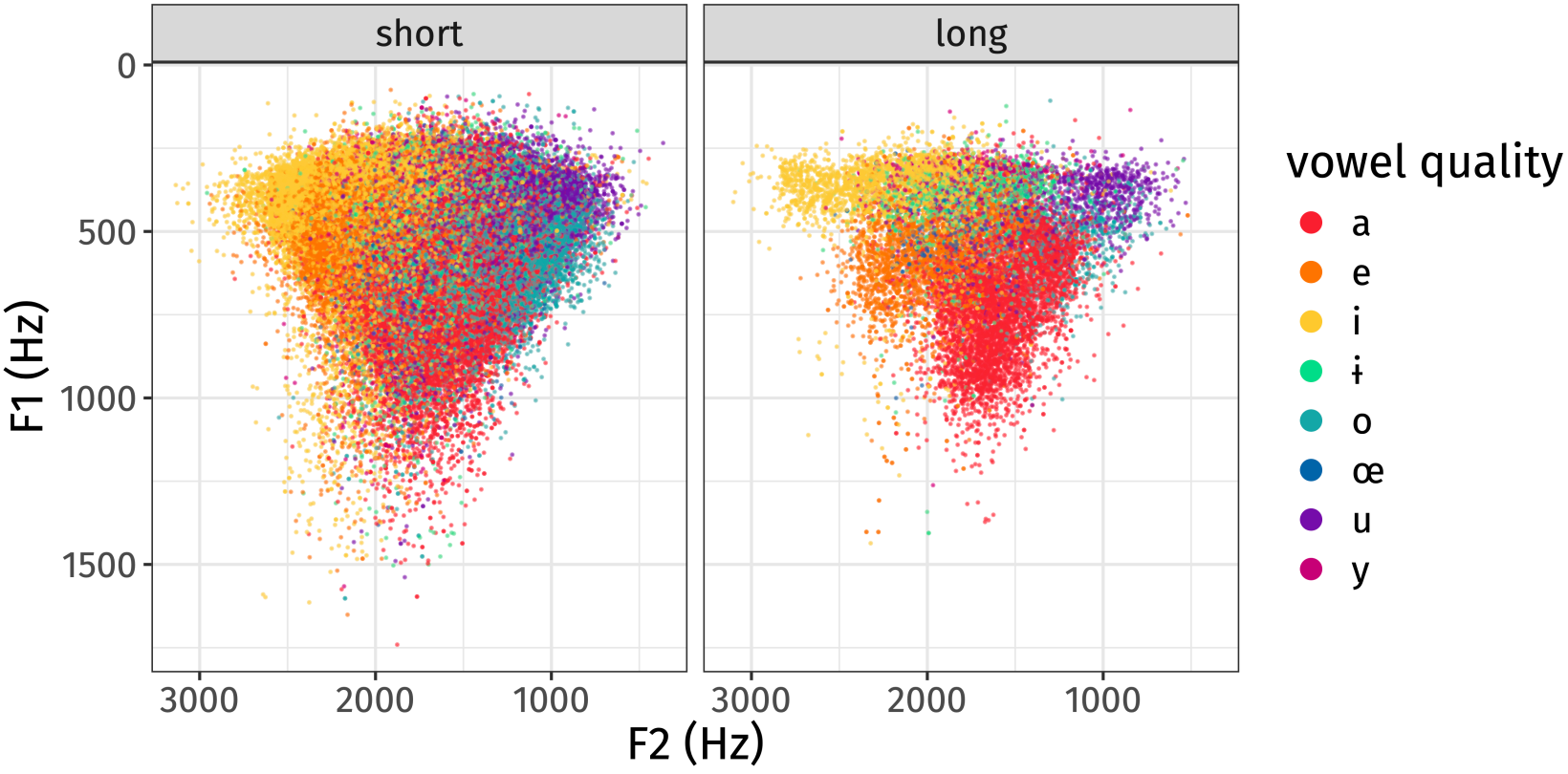
Data quality?



Uncorrected, *z*-score-normalised data.

Try. Automatically remeasure offenders with different ceilings.

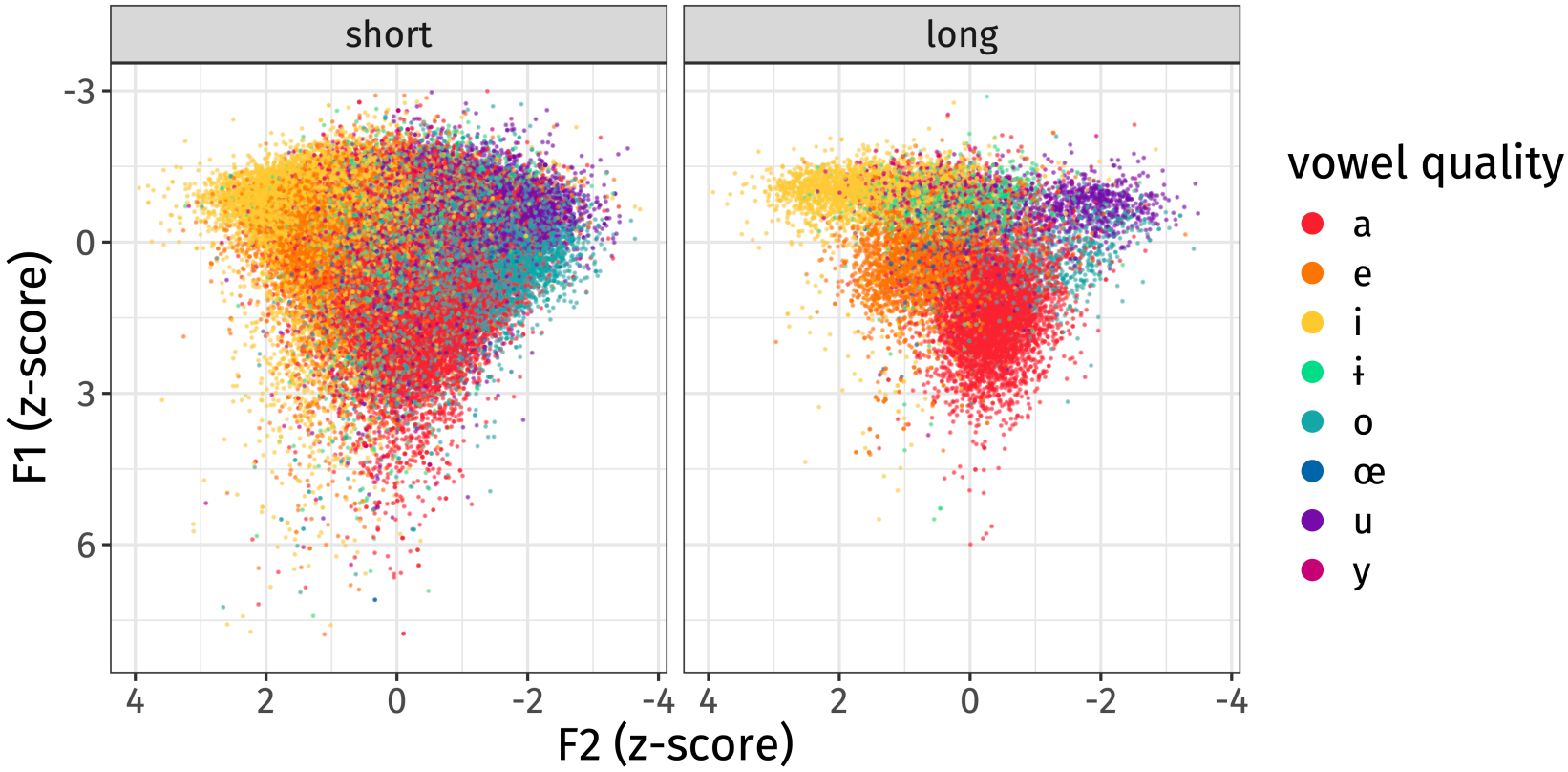
Data quality?



Revised, raw data.

Try. Automatically remeasure offenders with different ceilings.

Data quality?

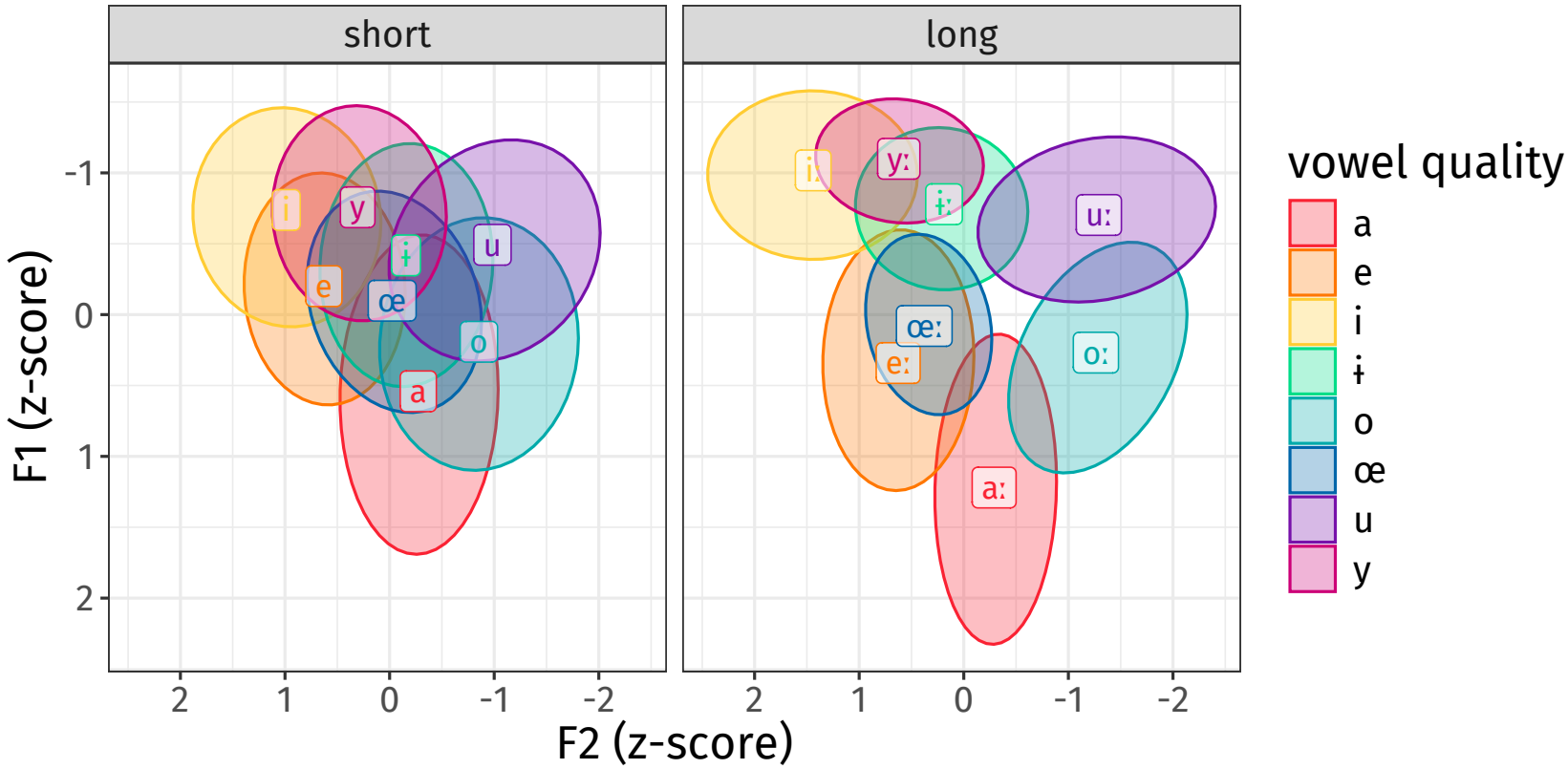


Revised, *z*-score-normalised data.

Result. Remeasured 30% of the data (by percentile) with adaptive ceiling between 4000–7000 Hz (number of formants = 5).

Properties of the vowel space

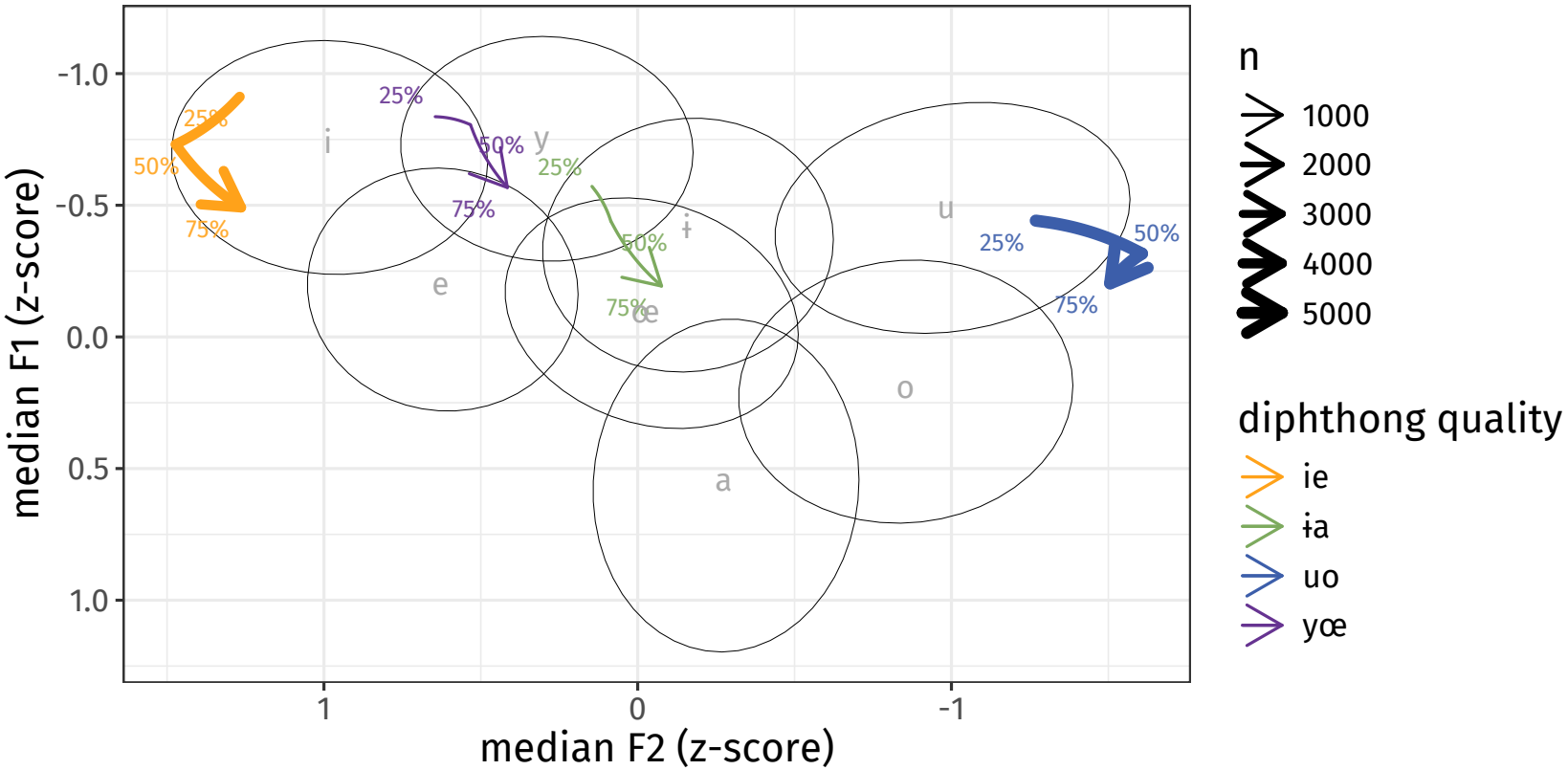
The overall picture



z-score normalised vowel space for monophthongs, 75% confidence.

Properties of the vowel space

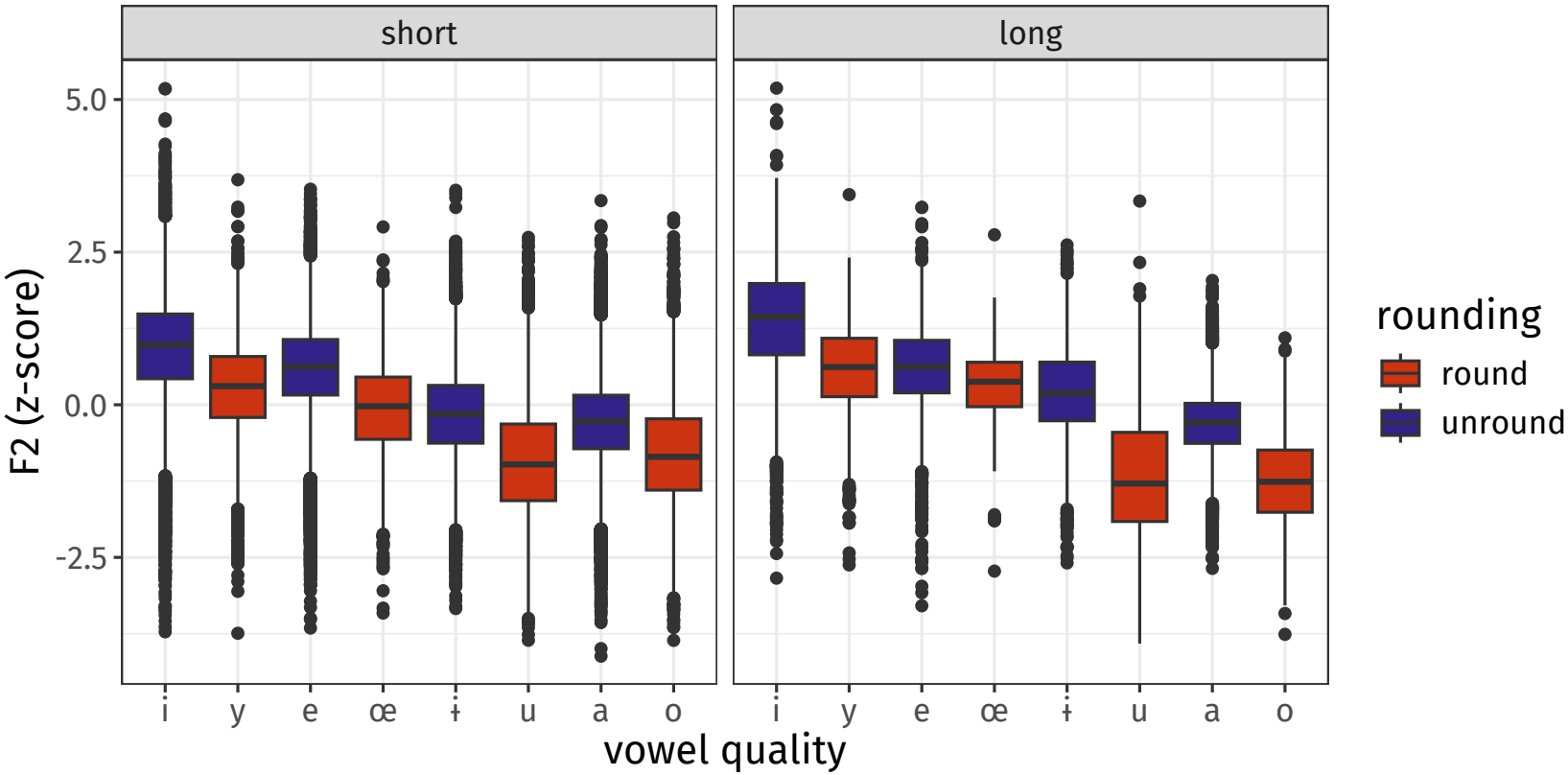
The diphthongs



25%, 50% & 75% means for (normalised) diphthong F1 and F2, shown with short monophthong CIs for reference.

Properties of the vowel space

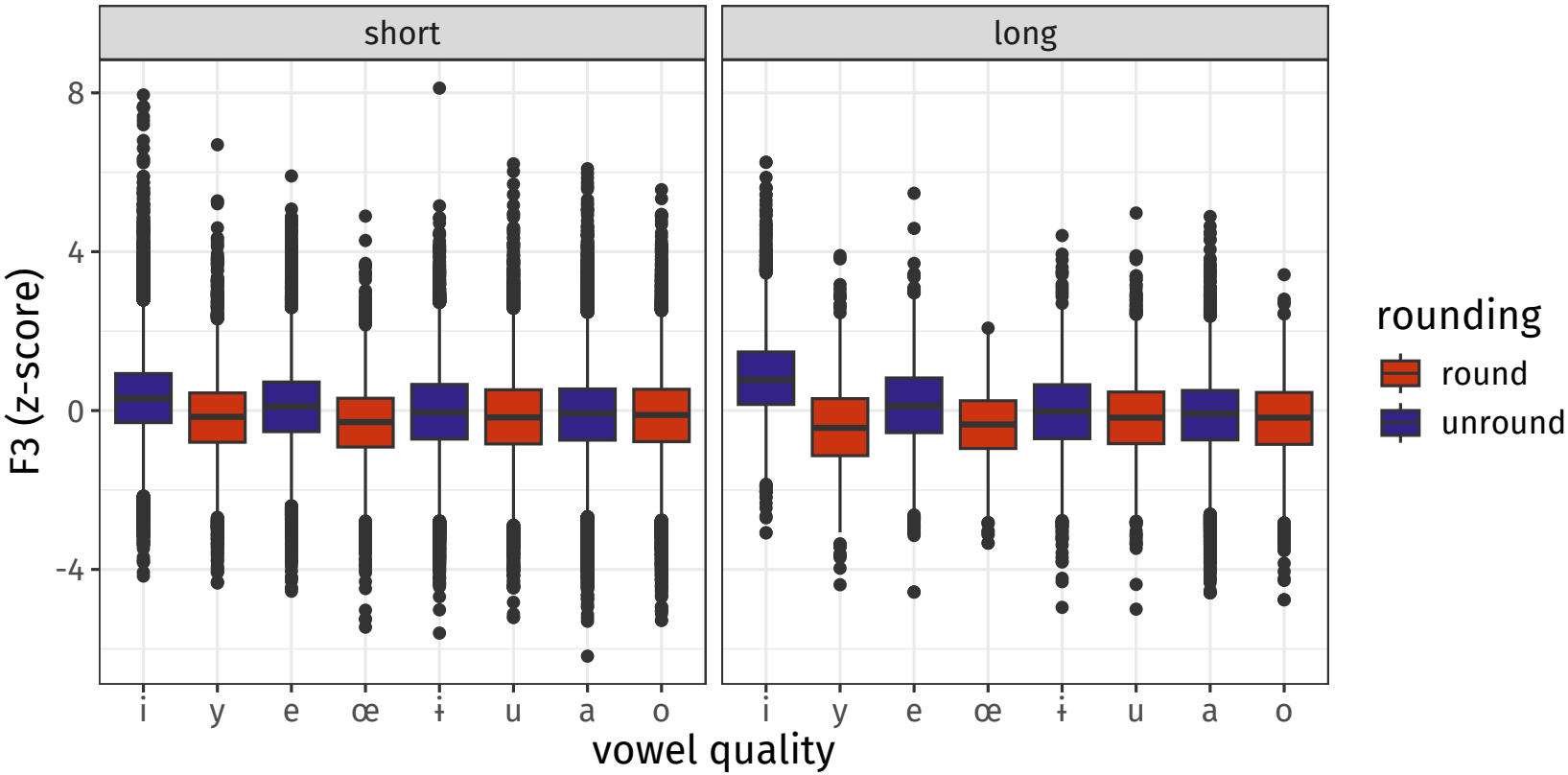
F2 & rounding



F2 better for distinguishing round-unround pairs than F3, as in Crimean Tatar & Kazakh (McCollum & Kavitskaya 2022; McCollum 2015).

Properties of the vowel space

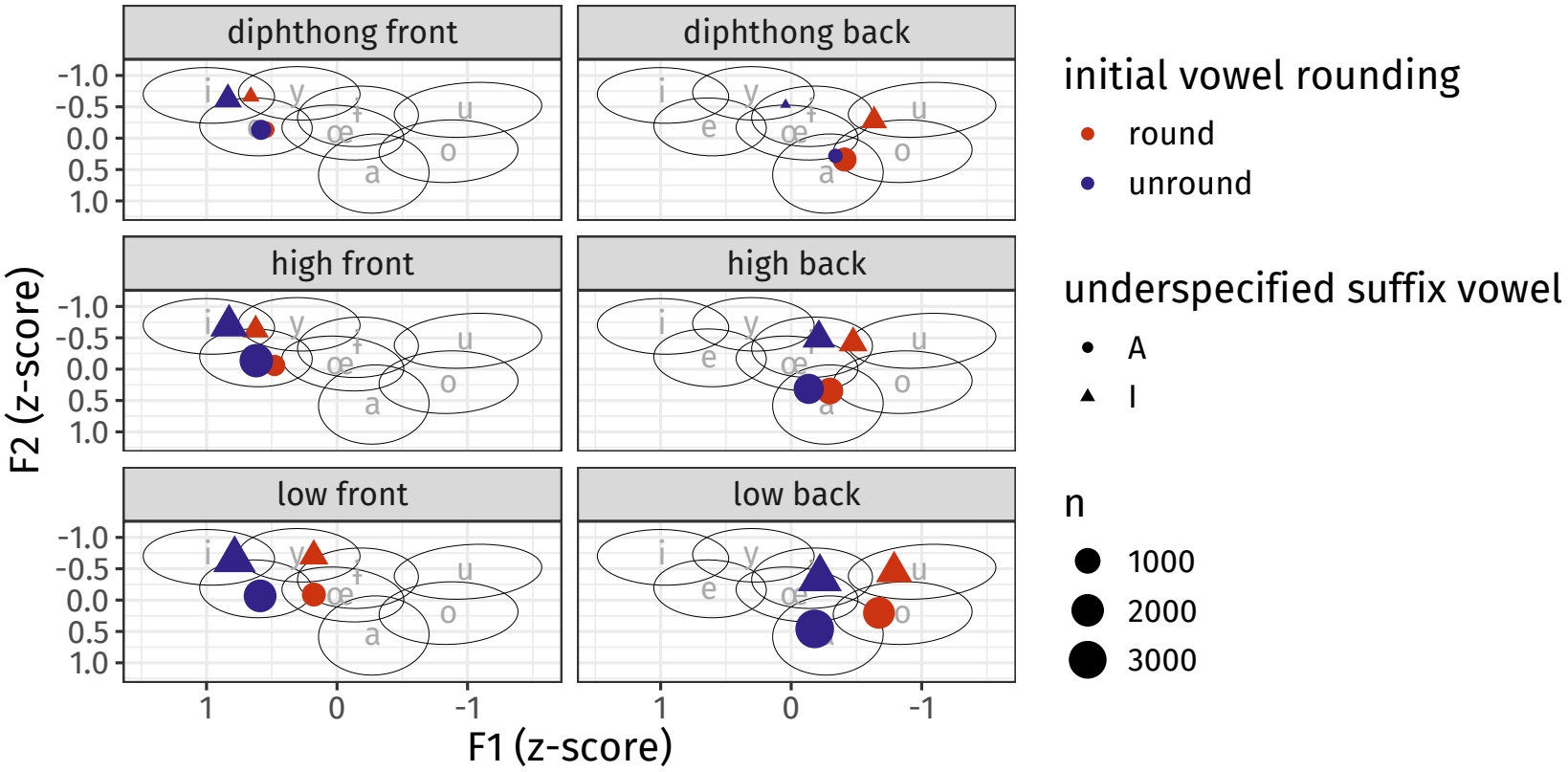
F3 & rounding



F2 better for distinguishing round-unround pairs than F3, as in Crimean Tatar & Kazakh (McCollum & Kavitskaya 2022; McCollum 2015).

Vowel harmony

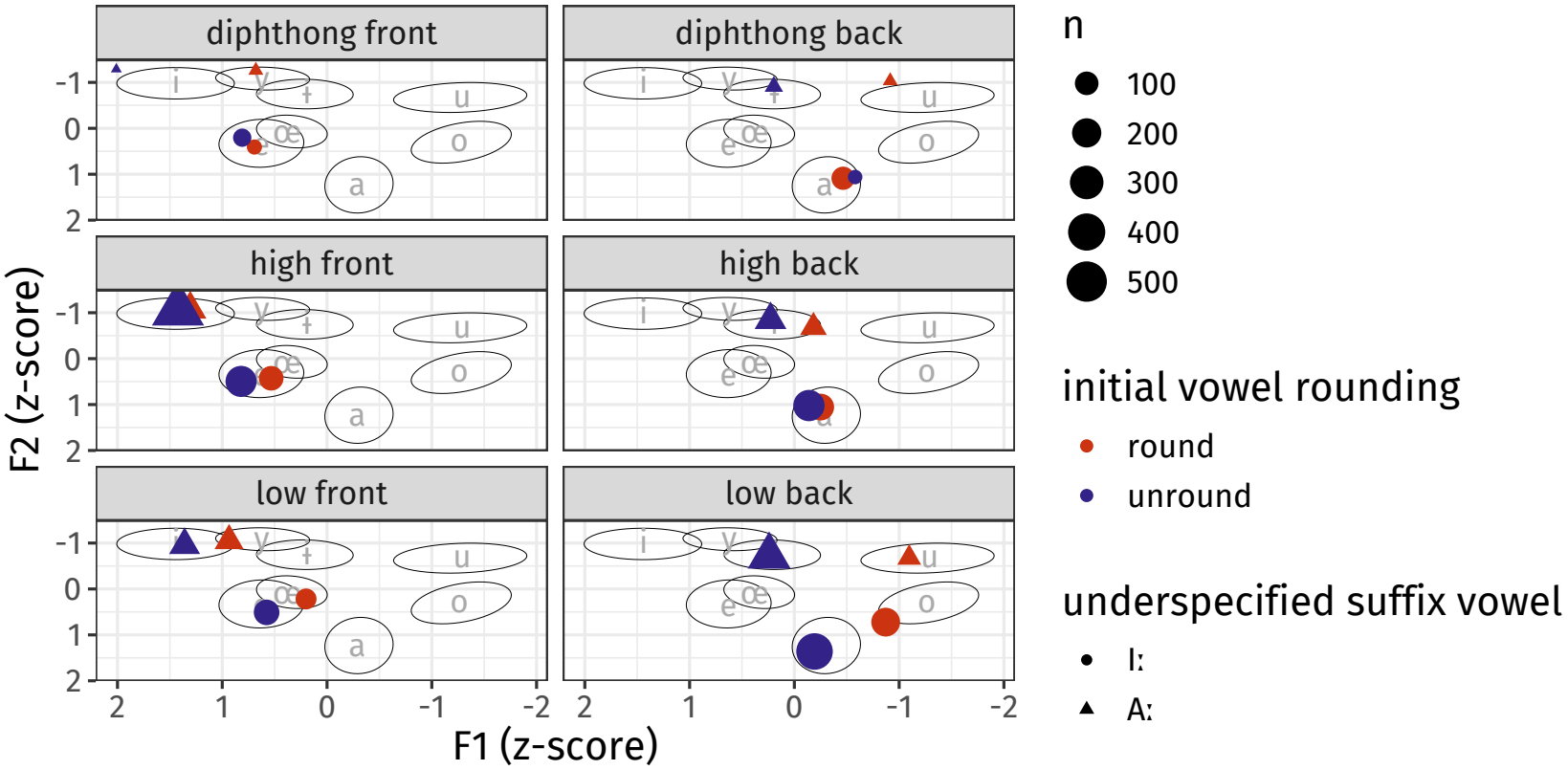
Underspecified short monophthongs in suffixes



Low vowels trigger rounding harmony across the board; **high vowels** are poorer triggers, and can only reliably trigger rounding harmony in high vowels. **Diphthongs pattern with high vowels as triggers of harmony.**

Vowel harmony

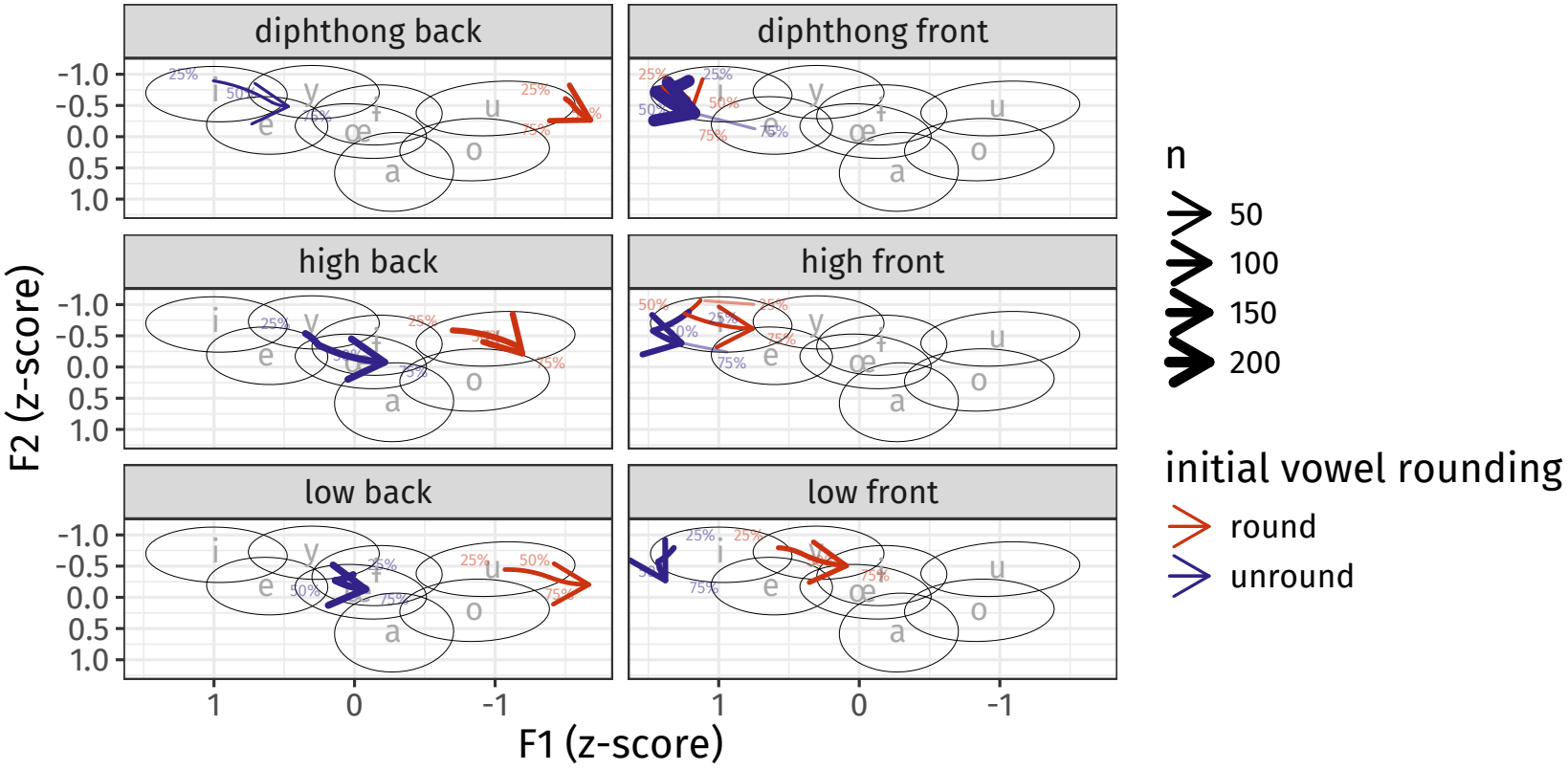
Underspecified long monophthongs in suffixes



Essentially similar patterning in the long vowels.

Vowel harmony

Underspecified diphthongs in suffixes

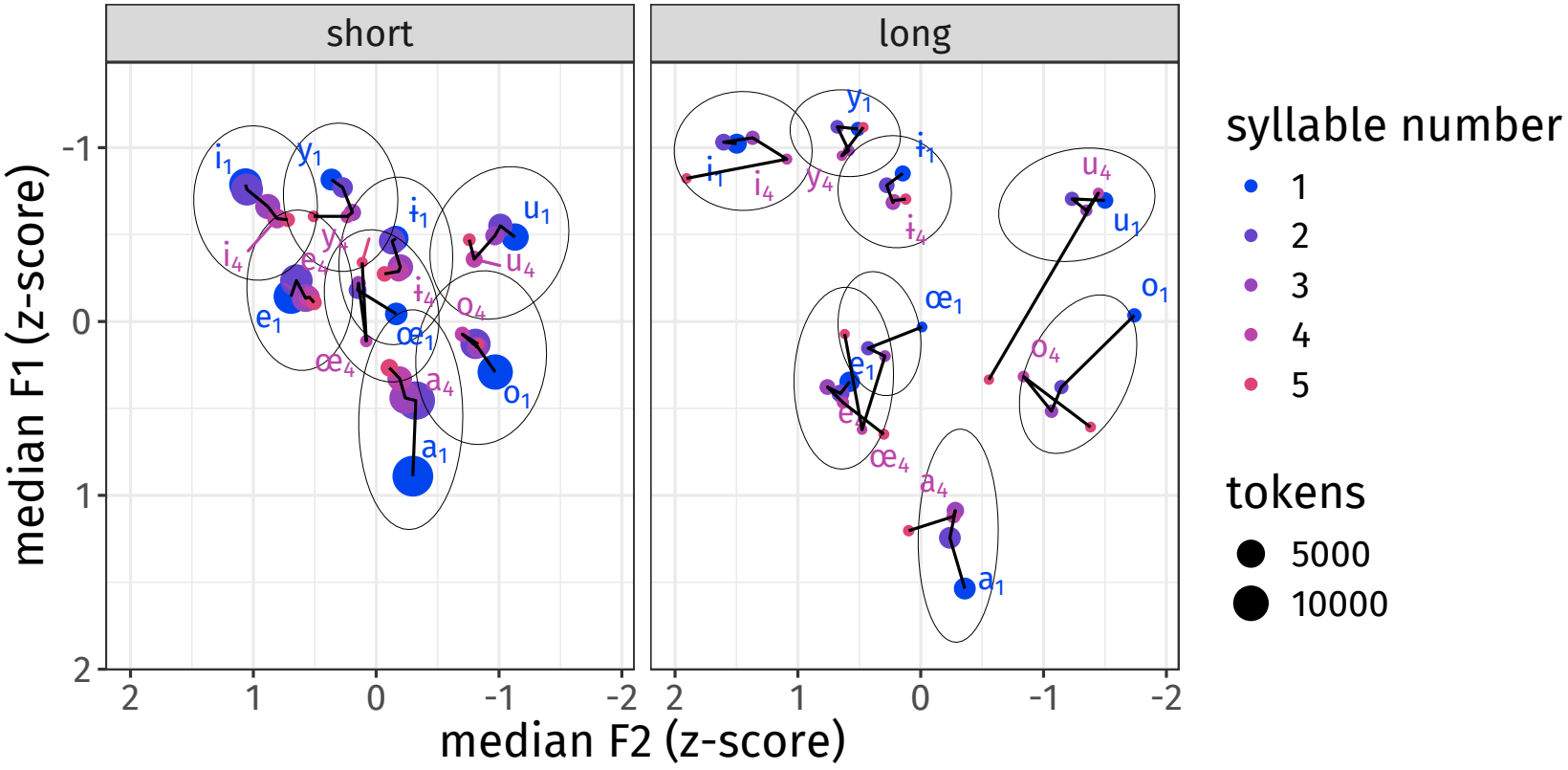


Affixes containing /IA/ underspecified diphthongs undergo harmony across the board (note small token numbers in post-/y/, post-/yœ̯/ position).

Diphthongs pattern with high vowels as targets, too.

Vowel harmony

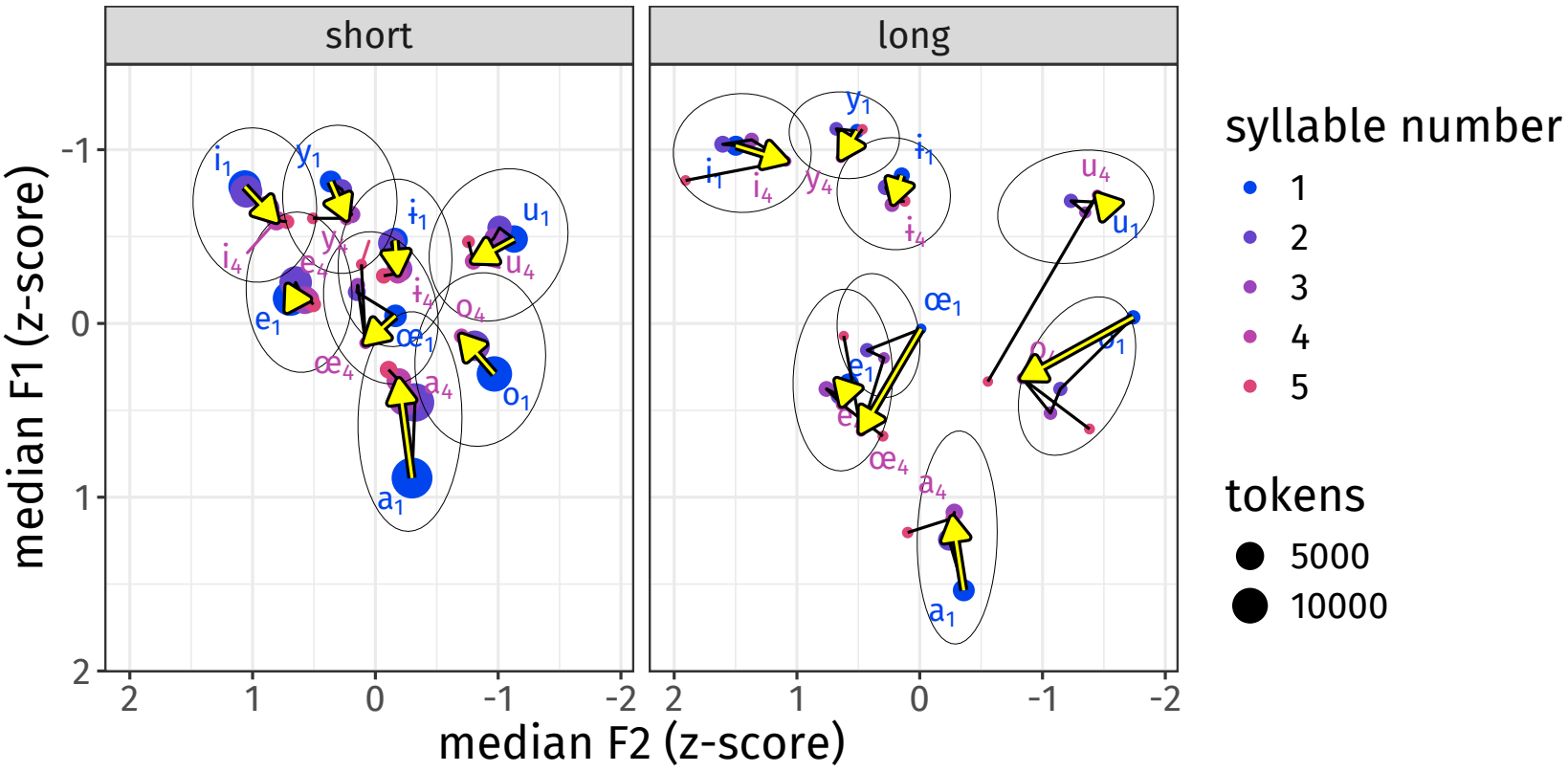
Centralisation & reduction



Short and long vowels organised by position in the word. **Short vowels centralise considerably** with distance from the initial syllable; no such systematic pattern for the long vowels.

Vowel harmony

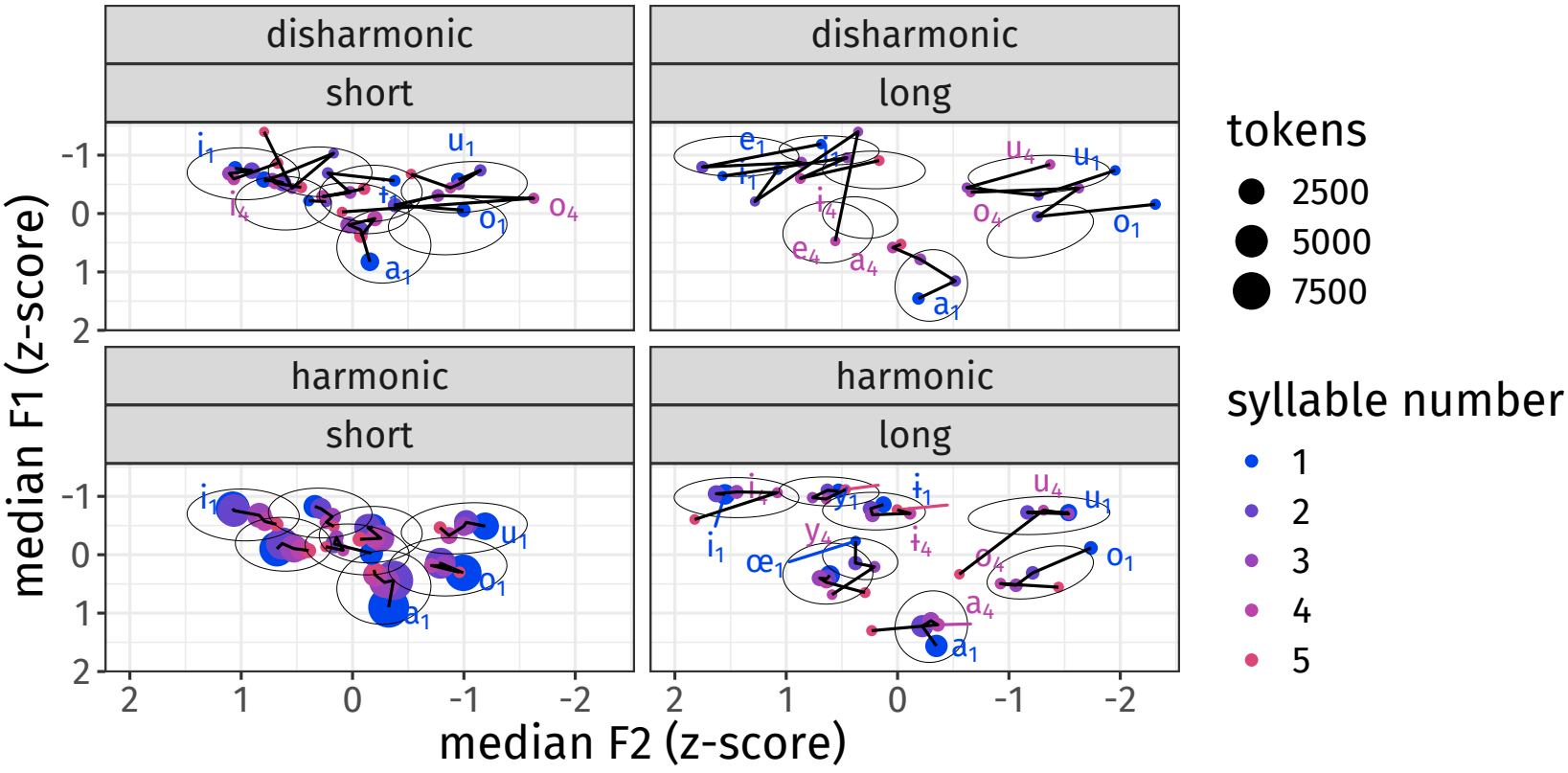
Centralisation & reduction



Short and long vowels organised by position in the word. **Short vowels centralise considerably** with distance from the initial syllable; no such systematic pattern for the long vowels.

Vowel harmony

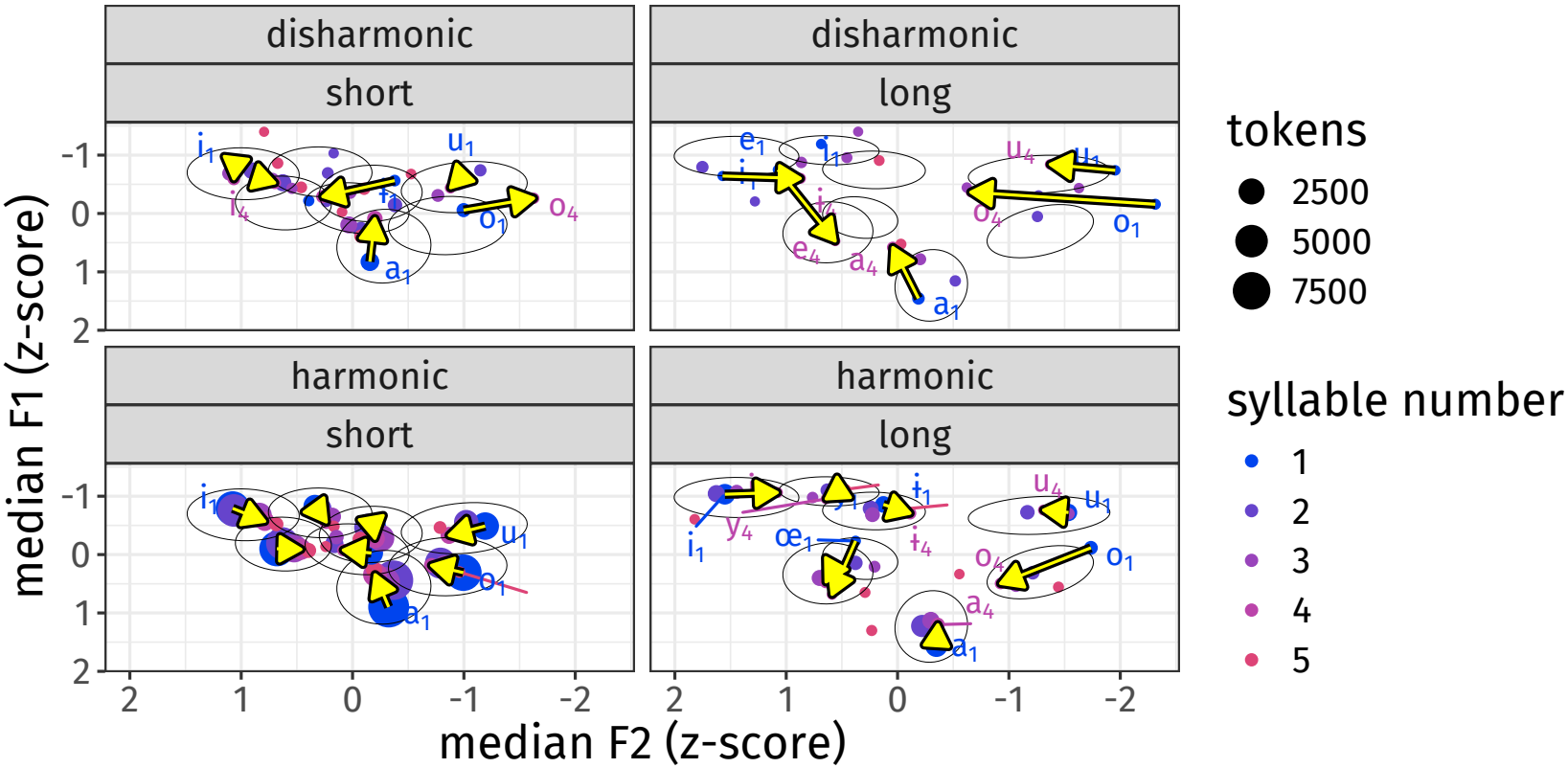
Centralisation & reduction



Centralisation is a property of harmonic items; **front-back disharmonic** words don't show it. **Centralisation = predictability.**

Vowel harmony

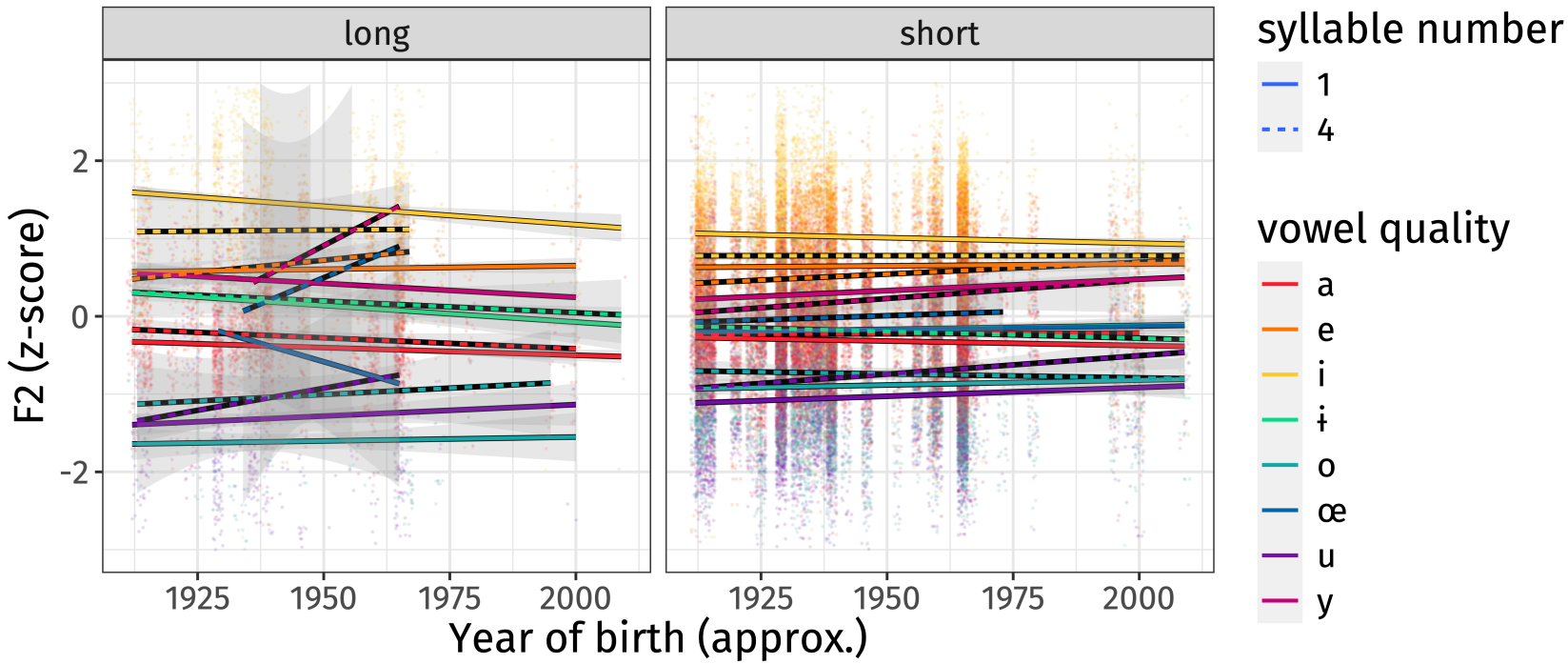
Centralisation & reduction



Centralisation is a property of harmonic items; **front-back disharmonic** words don't show it. **Centralisation = predictability.**

Vowel harmony

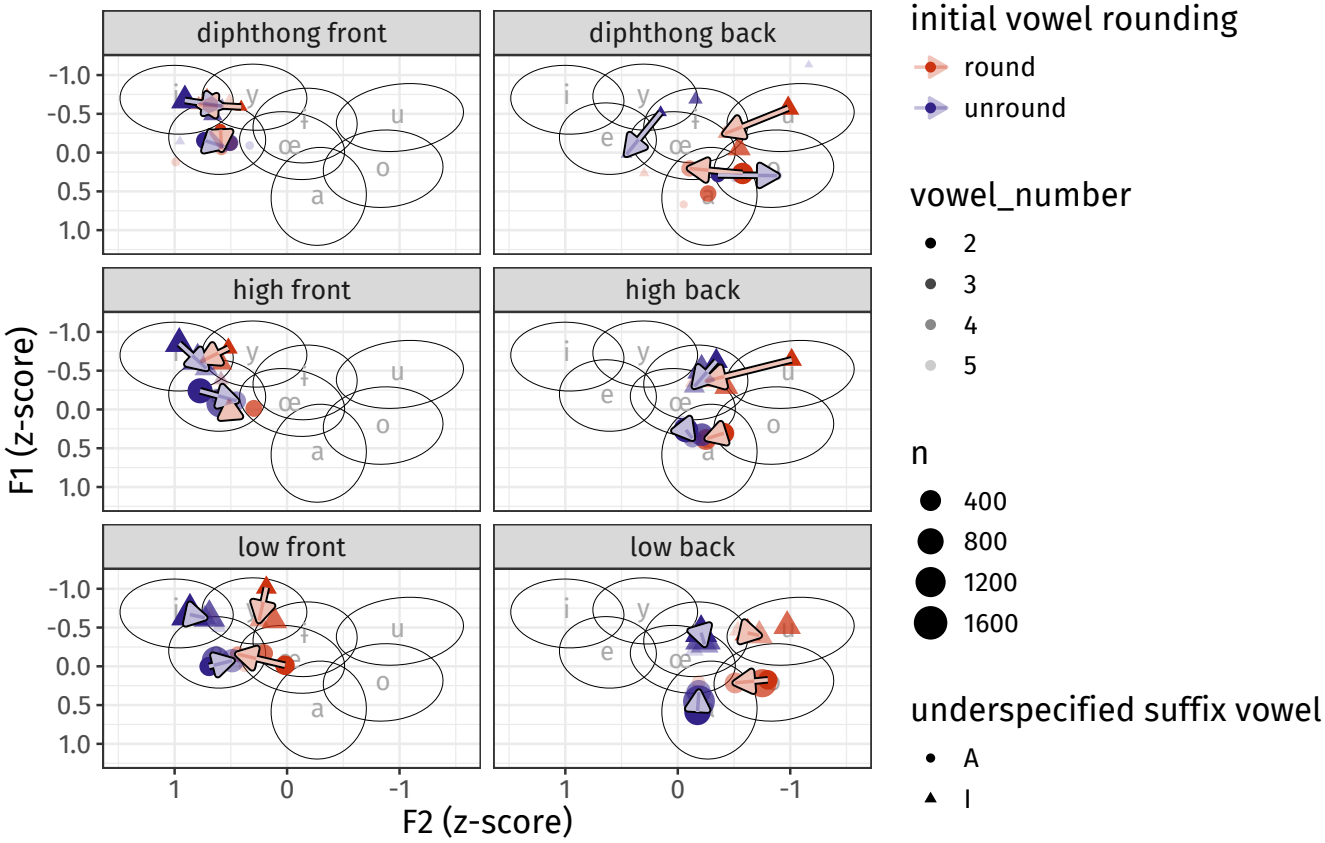
Centralisation & reduction



And no *divergence* between syllable-1 and syllable-4 short vowels in apparent time. **Centralisation is stable?**

Vowel harmony

Centralisation & reduction



So does this interact with VH? Distinguishability of A & I by backness and roundness of trigger in good shape until syllable 5, after which arguably driven more by dropoff in token numbers than by phonology. → **Not much.**

Summary & outlook

- **Dolgan is a boring, generic Turkic language with stable backness and rounding harmony.**

Summary & outlook

- **Dolgan is a boring, generic Turkic language with stable backness and rounding harmony.**
  - *Despite* being in a relatively extreme sociolinguistic condition ...

Summary & outlook

- **Dolgan is a boring, generic Turkic language with stable backness and rounding harmony.**
  - *Despite* being in a relatively extreme sociolinguistic condition …
  - *And* showing the predictability-driven centralisation that has often been suggested to be a precursor to harmony decay.

Summary & outlook

- **Dolgan is a boring, generic Turkic language with stable backness and rounding harmony.**
  - *Despite* being in a relatively extreme sociolinguistic condition ...
  - *And* showing the predictability-driven centralisation that has often been suggested to be a precursor to harmony decay.
- There is some *gradience*: backness and rounding do drop off as suffixes get further from the trigger, in line with reports from many other Turkic languages. But not enough to seriously threaten the system itself.

Summary & outlook

- **Dolgan is a boring, generic Turkic language with stable backness and rounding harmony.**
  - *Despite* being in a relatively extreme sociolinguistic condition ...
  - *And* showing the predictability-driven centralisation that has often been suggested to be a precursor to harmony decay.
- There is some *gradience*: backness and rounding do drop off as suffixes get further from the trigger, in line with reports from many other Turkic languages. But not enough to seriously threaten the system itself.
- There is plenty of *disharmony* in the lexicon, but it doesn't do anything.
- **Why does this work?**

Summary & outlook

- **Dolgan is a boring, generic Turkic language with stable backness and rounding harmony. Why?**

Summary & outlook

- **Dolgan is a boring, generic Turkic language with stable backness and rounding harmony.** **Why?**
- **Conjecture.** Phonemic length gives Dolgan a category of vowels which are:
    - Perceptually-salient

Summary & outlook

- **Dolgan is a boring, generic Turkic language with stable backness and rounding harmony.** **Why?**
- **Conjecture.** Phonemic length gives Dolgan a category of vowels which are:
  - Perceptually-salient
  - Don't undergo centralisation

Summary & outlook

- **Dolgan is a boring, generic Turkic language with stable backness and rounding harmony.** **Why?**
- **Conjecture.** Phonemic length gives Dolgan a category of vowels which are:
  - Perceptually-salient
  - Don't undergo centralisation
  - Are under-represented in the disharmonic (borrowed) component of the lexicon

Summary & outlook

- **Dolgan is a boring, generic Turkic language with stable backness and rounding harmony.** **Why?**
- **Conjecture.** Phonemic length gives Dolgan a category of vowels which are:
  - Perceptually-salient
  - Don't undergo centralisation
  - Are under-represented in the disharmonic (borrowed) component of the lexicon
  - Do harmonise

Summary & outlook

- **Dolgan is a boring, generic Turkic language with stable backness and rounding harmony.** **Why?**
- **Conjecture.** Phonemic length gives Dolgan a category of vowels which are:
  - Perceptually-salient
  - Don't undergo centralisation
  - Are under-represented in the disharmonic (borrowed) component of the lexicon
  - Do harmonise
- If true, ultimately, harmony in Dolgan is protected for reasons that perhaps also extend to lgs. like **Sakha** and **Kyrgyz**.

Summary & outlook

- **Dolgan is a boring, generic Turkic language with stable backness and rounding harmony.** **Why?**
- **Conjecture.** Phonemic length gives Dolgan a category of vowels which are:
  - Perceptually-salient
  - Don't undergo centralisation
  - Are under-represented in the disharmonic (borrowed) component of the lexicon
  - Do harmonise
- If true, ultimately, harmony in Dolgan is protected for reasons that perhaps also extend to lgs. like **Sakha** and **Kyrgyz**.
- **This is work in progress — further ideas very welcome!**

[pasʲibalaːt͡ʃːɨɬar]!

Supported by Riksbankens Jubileumsfond grant P23-0791 (2024–2027)
‘The trajectory and distributional typology of phonological change’.

References I

Binnick, Robert I. 1991. Vowel harmony loss in Uralic and Altaic. In William G. Boltz & Michael C. Shapiro (eds.), *Studies in the historical phonology of Asian languages* (Current issues in linguistic theory 77), 35–52. Amsterdam: John Benjamins.

Bobaljik, Jonathan David. 2018. Disharmony and decay: Itelmen vowel harmony in the 20th century. In Roberto Petrosino, Pietro Cerrone & Harry van der Hulst (eds.), *From sounds to structures: Beyond the Veil of Maya*, 161–192. Berlin: Mouton de Gruyter.

Chan, May Pik Yu & Jianjing Kuang. 2023. The acoustics of vowel harmony in Sakha. *ICPhS 2023: Phonetics of Lesser Documented and Endangered Languages* 19(331). 3296–3300.

Conklin, Jenna & Olga Dmitrieva. 2018. Acoustics of Tatar vowels: Articulation and vowel-to-vowel coarticulation. *The Journal of the Acoustical Society of America* 144(3). 1940. doi:10.1121/1.5068486.

Dawkins, Richard McGillivray & William Reginald Halliday. 1916. *Modern Greek in Asia Minor a study of the dialects of Síli, Cappadocia and Phárasa, with grammar, texts, translations and glossary*. Cambridge: Cambridge University Press.

Däbritz, Chris Lasse. 2020. The INEL Dolgan corpus: Insights into an endangered language of Northern Eurasia. *Finno-Ugric Languages and Linguistics* 9(1–2). 67–77. http://full.btk.ppke.hu.

Däbritz, Chris Lasse. 2022. *A grammar of Dolgan: A northern Siberian Turkic language of the Taimyr peninsula*. Leiden: Brill. doi:10.1163/9789004516427.

Däbritz, Chris Lasse, Nina Kudryakova & Eugénie Stapert. 2022. INEL Dolgan corpus. version 2.0. Archived at University of Hamburg. Publication date 2022-11-30. https://hdl.handle.net/11022/0000-0007-F9A7-4.

Harrison, K. David, Mark Dras & Berk Kapicioglu. 2006. Agent-based modeling of the evolution of vowel harmony. In M. Hirotani (ed.), *Proceedings of the Northeast Linguistic Society*, vol. 32, 217–236.

Hyman, Larry M. 2002. Is There a Right-to-Left Bias in Vowel Harmony? Talk given at the 9th International Phonology Meeting, Vienna, Austria, 1 November. URL: http://www.linguistics.berkeley.edu/~hyman/Hyman_Vienna_VH_paper_forma.pdf.

Kavitskaya, Darya. 2013. Segmental inventory and the evolution of harmony in Crimean Tatar. *Turkic Languages* 17. 86–114.

Kavitskaya, Darya & Adam McCollum. 2023. The rise and fall of rounding harmony in Turkic. In Darya Kavitskaya & Alan C. L. Yu (eds.), *The life cycle of language*, Oxford: Oxford University Press. doi:10.1093/oso/9780192845818.003.0002.

Li, Bing. 1996. *Tungusic vowel harmony: Description and analysis*: University of Amsterdam PhD thesis.

McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner & Morgan Sonderegger. 2017. Montreal Forced Aligner: trainable text-speech alignment using Kaldi. In *Proceedings of the 18th conference of the international speech communication association*, .

References II

McCollum, Adam G. 2015. Labial harmonic shift in Kazakh: Mapping the pathways and motivations for decay. In Anna E. Jurgensen, Hannah Sande, Spencer Lamoureux, Kenny Baclawski & Allison Zerbe (eds.), *The proceedings of the 41st annual meeting of the berkeley linguistics society*, 329–352. doi:10.20354/B4414110012.

McCollum, Adam G. 2020. Vowel harmony and positional variation in Kyrgyz. *Laboratory Phonology* 11(1).

McCollum, Adam G. 2024. On how and why vowel harmony decays. In Nancy A. Ritter & Harry van der Hulst (eds.), *The oxford handbook of vowel harmony*, 565–573. Oxford: Oxford University Press.

McCollum, Adam G. & Si Chen. 2020. Kazakh. *Journal of the International Phonetic Association* 51(2). 276–298. doi: 10.1017/s0025100319000185.

McCollum, Adam G., Karthik Durvasula & Xiayimaierdan Abudushalamu. 2024. O gradience, where art thou? Examining backness harmony in Uyghur. Rutgers University and Michigan State University, ms. URL: https://ling.auf.net/lingbuzz/007939.

McCollum, Adam G & Darya Kavitskaya. 2022. On the status of non-iterativity in feature spreading. *Glossa: a journal of general linguistics* 7(1). doi:10.16995/glossa.5783. http://dx.doi.org/10.16995/glossa.5783.

Pearce, Mary. 2008. Vowel harmony domains and vowel undershoot. *UCL Working Papers in Linguistics* 20. 115–140.

Pearce, Mary. 2012. Effects of harmony on reduction in Kera. *Linguistic Variation* 12(2). 292–320.

Sandstedt, Jade J. 2020. Vowel harmony decay in Old Norwegian. *Papers in Historical Phonology* 5. 11–48.

Shiraishi, Hidetoshi & Bert Botma. 2017. On the diachronic origin of Nivkh height restrictions. In Geoff Lindsey & Andrew Nevins (eds.), *Sonic signatures*, 201–214. Amsterdam: John Benjamins.

Sjoberg, Andrée F. 1963. *Uzbek structural grammar*. Bloomington, Indiana: Indiana University Press.