Error detection and correction in L1 and L2

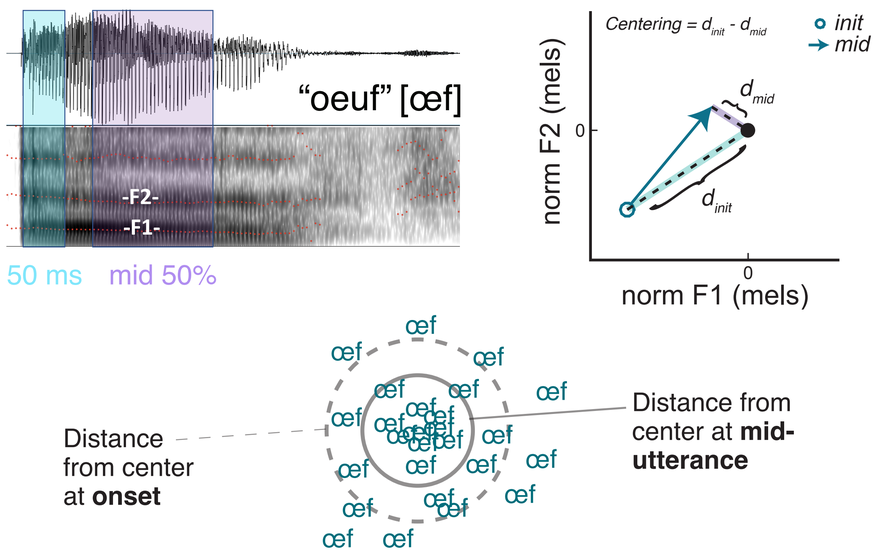
Sarah Bakst and Caroline A. Niziolek

University of Wisconsin–Madison, Dept. of Communication Sciences and Disorders + Waisman Center

How Do We Use Auditory Feedback When Speaking a Non-native Language?

When we speak in our native language (L1), there is strong evidence that we use the auditory feedback of our speech to monitor and correct for errors:

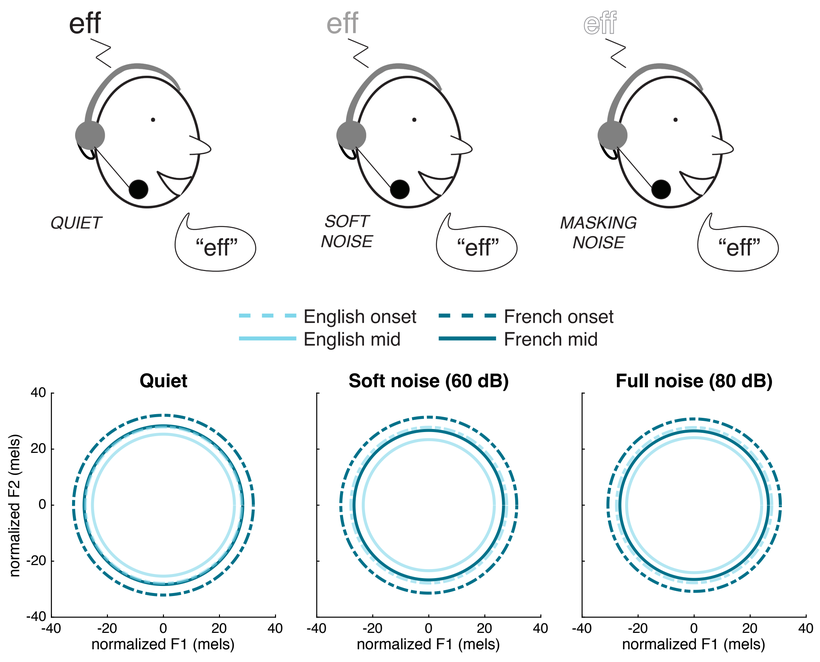
- Alterations to auditory feedback (e.g. when speakers hear themselves producing speech with an imposed formant shift) cause an online compensatory change to motor plans that results in reduced auditory error.[1-3]
- Vowel formant trajectories reflect a "**centering**" behavior: individual productions converge on values that represent their median productions, reducing variability over the course of an utterance[4] (e.g. from the blue window to the purple window, below).
- Masking auditory feedback with noise can reduce centering[5], suggesting centering is partially driven by auditory feedback.



How do we use auditory feedback in a less familiar language, such as a second language (L2) learned in adulthood? In three experiments, we investigate auditory feedback control in L2 speakers of French. All participants were native English speakers with self-reported normal speech and hearing who began learning French in adulthood (minimum of two semesters of French instruction).

Formant Variability and Vowel Centering

The noise masking experiment investigated how auditory feedback affects variability and formant dynamics in L2. Participants (n = 12) read aloud 40 repetitions of each stimulus word in English and French in each of three noise levels: quiet, soft noise, and full noise, which was calibrated to fully mask auditory feedback. A visual loudness target was provided on each trial to avoid involuntary changes in loudness from the Lombard effect.



English (L1) stimuli: *Eve*, *eff*, *add* [iv] [ɛf] [æd]

French (L2) stimuli: *Yves*, *hais*, *oeuf* [iv] [ɛ/e] [œf]

We measured trial-to-trial variability in a time window at utterance onset (first 50 ms of each trial; dashed lines) and a mid-utterance time window (middle 50% of each trial; solid lines). **Variability** was defined as the average F1-F2 (mel) distance to the median production defined separately in each time window. **Centering** was defined as the difference between these two distances; i.e., the reduction in variability from the initial to the middle time window.

Variability:

- Variability was greater in French than in English in both onset and mid-utterance time windows ($p < 0.009$, $p < 0.04$).
  - Better feedforward control in L1 may enable more precise speech movements, resulting in smaller variability in L1 than in L2.
- Variability did not significantly differ by noise level.
  - In contrast to previous findings, there was no tendency to increase variability as auditory feedback was increasingly masked by noise.

Centering:

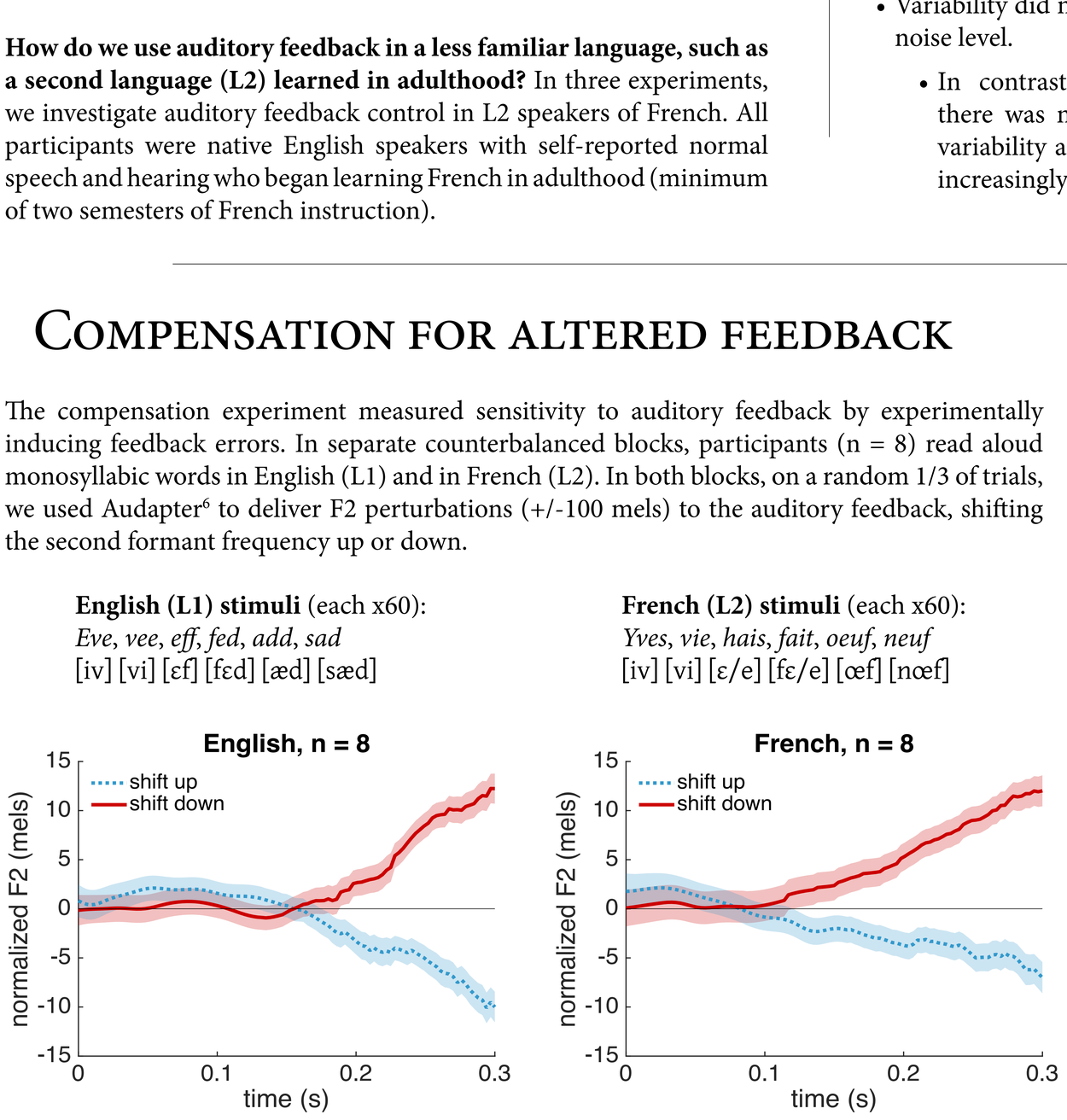
- Both languages showed evidence of centering: a reduction in variability from vowel onset to vowel midpoint.
- Centering was, on average, *greater* in French than in English in quiet (5.5 mels) and soft noise (8.5 mels), but it did not differ between languages in full masking noise.
  - A greater reliance on feedback control in L2 could result in greater within-utterance reduction of formant variability (i.e., centering) than in L1 (even as this greater centering fails to overcome the greater initial variability). This is consistent with no difference between the languages in full noise, when auditory feedback was not available.
- However, in contrast to previous findings, centering did not significantly differ between full noise and quiet (4 mels, $p = 0.12$), but was greater in soft noise than quiet (7 mels, $p = 0.02$).

Compensation for altered feedback

The compensation experiment measured sensitivity to auditory feedback by experimentally inducing feedback errors. In separate counterbalanced blocks, participants (n = 8) read aloud monosyllabic words in English (L1) and in French (L2). In both blocks, on a random 1/3 of trials, we used Audapter[6] to deliver F2 perturbations (+/-100 mels) to the auditory feedback, shifting the second formant frequency up or down.

English (L1) stimuli (each x60):
Eve, *vee*, *eff*, *fed*, *add*, *sad*
[iv] [vi] [ɛf] [fɛd] [æd] [sæd]

French (L2) stimuli (each x60):
Yves, *vie*, *hais*, *fait*, *oeuf*, *neuf*
[iv] [vi] [ɛ/e] [fɛ/e] [œf] [nœf]



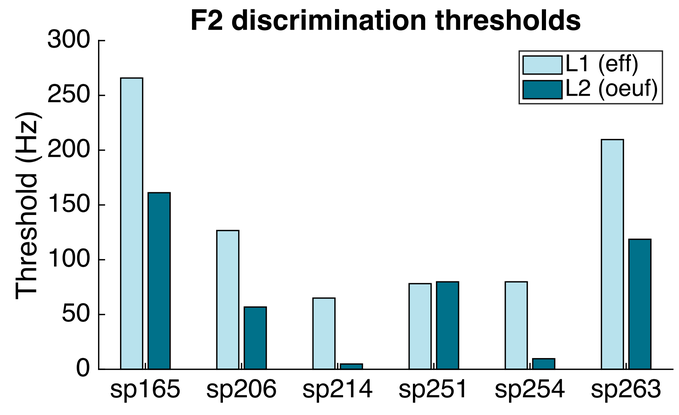
Formants were tracked over the course of the syllable using wave_viewer[7], a Matlab interface to Praat.[8] F2 trajectories from shift up and shift down trials were normalized by subtracting the time-aligned F2 trajectories from trials with normal feedback.

Preliminary results show similar magnitudes of compensation in both languages, but an earlier compensation onset in French than in English (Monte Carlo simulations, 1000x).

Self-perception in L1 and L2

Differences in feedback control between L1 and L2 may be a result of differences in sensitivity to fine-grained acoustic change in the two languages. To test this hypothesis, the participants in the compensation study also took part in a perception experiment to determine their sensitivity to formant changes in their own productions of English (L1) and French (L2) vowels.

- **Stimuli:** Altered versions of participants' own productions of "eff" and "oeuf" in which F2 had been shifted down in increments of 1 Hz.
- **Task:** Participants (n = 7) discriminated between formant-shifted and unshifted versions of their productions an AAXA paradigm (AABA or ABAA presentation), reporting whether the second or third stimulus was different from the others.



Preliminary results suggest greater sensitivity to subphonemic formant differences in French than English: that is, the L2 vowel had lower discrimination thresholds (shown here in Hz but also significantly lower when converted to mels).

Conclusions (TL;DR)

- Production variability was greater in L2 than in L1, reflecting less consistent feedforward control.
- This greater variability at onset in L2 was accompanied by a greater corrective movement (centering); however, variability remained greater in L2 even after the self-correction (i.e., centering is not necessarily indicative of skill).
- Compensation onset was earlier in French relative to English, possibly reflecting a greater reliance on auditory feedback in L2.
- Discrimination thresholds were lower in L2, possibly reflecting less categorical vowel perception[9] near L2 categories that don't exist in L1.

References

1. Niziolek, C.A., and Guenther, F.H. (2013). Vowel Category Boundaries Enhance Cortical and Behavioral Responses to Speech Feedback Alterations. J. Neurosci. 33, 12090–12098.
2. Purcell, D.W., and Munhall, K.G. (2006). Compensation following real-time manipulation of formants in isolated vowels. J. Acoust. Soc. Am. 119, 2288–2297.
3. Tourville, J.A., Reilly, K.J., and Guenther, F.H. (2008). Neural mechanisms underlying auditory feedback control of speech. NeuroImage 39, 1429–1443.
4. Niziolek, C.A., Nagarajan, S.S., and Houde, J.F. (2013). What Does Motor Efference Copy Represent? Evidence from Speech Production. J. Neurosci. 33, 16110–16116.
5. Niziolek, C.A., Nagarajan, S.S., and Houde, J.F. (2015). The contribution of auditory feedback to corrective movements in vowel formant trajectories. In The Scottish Consortium for ICPhS 2015 (Ed.), Proceedings of the 18th International Congress of Phonetic Sciences (The University of Glasgow).
6. Cai, S., Boucek, M., Ghosh, S.S., Guenther, F.H., and Perkell, J.S. (2008). A system for online dynamic perturbation of formant trajectories and results from perturbations of the Mandarin triphthong /iau/. Proc. 8th Intl Semin. Speech Prod., 65–68.
7. Niziolek, C.A., and Houde, J.F. (2015). wave_viewer: First release.
8. Boersma, P., and Weenink, D. (2009). Praat: doing phonetics by computer (Version 5.1.05).
9. Kuhl, P.K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. Percept. Psychophys. 50, 93–107.

Brain, Language, and Acoustic Behavior Lab

Waisman Center