What is Articulatory Control?

Structural Complexity in Natural Language: Paris

John Kingston

Department of Linguistics, University of Massachusetts, Amherst
jkingston@linguist.umass.edu

May 31, 2016

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Jakobson, Fant, & Halle (1952: 13)

"We speak to be heard in order to be understood"

Outline of the talk

1. The original proposal (Kingston & Diehl, 1994),
2. New data regarding [voice] and F0 perturbations,
3. A puzzle: Time and F0 change,
4. Do speakers care about listeners?
5. Coarticulation, sound change, and the distribution of information in the signal.

Outline of the talk

1. The original proposal (Kingston & Diehl, 1994),
2. New data regarding [voice] and F0 perturbations,
3. A puzzle: Time and F0 change,
4. Do speakers care about listeners?
5. Coarticulation, sound change, and the distribution of information in the signal.

Outline of the talk

1. The original proposal (Kingston & Diehl, 1994),
2. New data regarding [voice] and F0 perturbations,
3. A puzzle: Time and F0 change,
4. Do speakers care about listeners?
5. Coarticulation, sound change, and the distribution of information in the signal.

Outline of the talk

1. The original proposal (Kingston & Diehl, 1994),
2. New data regarding [voice] and F0 perturbations,
3. A puzzle: Time and F0 change,
4. Do speakers care about listeners?
5. Coarticulation, sound change, and the distribution of information in the signal.

Outline of the talk

1. The original proposal (Kingston & Diehl, 1994),
2. New data regarding [voice] and F0 perturbations,
3. A puzzle: Time and F0 change,
4. Do speakers care about listeners?
5. Coarticulation, sound change, and the distribution of information in the signal.

The original proposal (Kingston & Diehl, 1994)

1. Observation: Each distinctive feature has many acoustic correlates, which vary between contexts and languages,
2. Example. [voice]. Voicing/voice onset time and F0 in following Vs:
   a [+voice]: voiced or earlier onset and lower F0,
   b [-voice]: voiceless or later onset and higher F0;
3. Assumption: Voicing or its timing is controlled, and F0 covaries,
4. Question: What's the mechanism responsible for F0 covariation?
5. Suggestions:
   a Cricothyroid contraction in [-voice] Cs to suppress voicing carries over into following Vs, raising their F0 (Löfqvist, et al., 1989; see also Halle & Stevens's, 1971, proposal for [stiff] versus [slack] folds),
   b Larynx lowering in [+voice] Cs to expand the oral cavity and slow the oral air pressure rise also lowers F0 (Hombert, et al., 1979; see Honda, et al. (1999): vocal folds are shortened by larynx lowering).

The original proposal (Kingston & Diehl, 1994)

1. Observation: Each distinctive feature has many acoustic correlates, which vary between contexts and languages,
2. Example. [voice]. Voicing/voice onset time and F0 in following Vs:
   a [+voice]: voiced or earlier onset and lower F0,
   b [-voice]: voiceless or later onset and higher F0;
3. Assumption: Voicing or its timing is controlled, and F0 covaries,
4. Question: What's the mechanism responsible for F0 covariation?
5. Suggestions:
   a Cricothyroid contraction in [-voice] Cs to suppress voicing carries over into following Vs, raising their F0 (Löfqvist, et al., 1989; see also Halle & Stevens's, 1971, proposal for [stiff] versus [slack] folds),
   b Larynx lowering in [+voice] Cs to expand the oral cavity and slow the oral air pressure rise also lowers F0 (Hombert, et al., 1979; see Honda, et al. (1999): vocal folds are shortened by larynx lowering).

The original proposal (Kingston & Diehl, 1994)

1. Observation: Each distinctive feature has many acoustic correlates, which vary between contexts and languages,
2. Example. [voice]. Voicing/voice onset time and F0 in following Vs:
   a. [+voice]: voiced or earlier onset and lower F0,
   b. [-voice]: voiceless or later onset and higher F0;
3. Assumption: Voicing or its timing is controlled, and F0 covaries,
4. Question: What's the mechanism responsible for F0 covariation?
5. Suggestions:
   a. Cricothyroid contraction in [-voice] Cs to suppress voicing carries over into following Vs, raising their F0 (Löfqvist, et al., 1989; see also Halle & Stevens's, 1971, proposal for [stiff] versus [slack] folds),
   b. Larynx lowering in [+voice] Cs to expand the oral cavity and slow the oral air pressure rise also lowers F0 (Hombert, et al., 1979; see Honda, et al. (1999): vocal folds are shortened by larynx lowering).

The original proposal (Kingston & Diehl, 1994)

1. Observation: Each distinctive feature has many acoustic correlates, which vary between contexts and languages,
2. Example. [voice]. Voicing/voice onset time and F0 in following Vs:
   a [+voice]: voiced or earlier onset and lower F0,
   b [-voice]: voiceless or later onset and higher F0;
3. Assumption: Voicing or its timing is controlled, and F0 covaries,
4. Question: What's the mechanism responsible for F0 covariation?
5. Suggestions:
   a Cricothyroid contraction in [-voice] Cs to suppress voicing carries over into following Vs, raising their F0 (Löfqvist, et al., 1989; see also Halle & Stevens's, 1971, proposal for [stiff] versus [slack] folds),
   b Larynx lowering in [+voice] Cs to expand the oral cavity and slow the oral air pressure rise also lowers F0 (Hombert, et al., 1979; see Honda, et al. (1999): vocal folds are shortened by larynx lowering).

The original proposal (Kingston & Diehl, 1994)

1. Observation: Each distinctive feature has many acoustic correlates, which vary between contexts and languages,
2. Example. [voice]. Voicing/voice onset time and F0 in following Vs:
   a [+voice]: voiced or earlier onset and lower F0,
   b [-voice]: voiceless or later onset and higher F0;
3. Assumption: Voicing or its timing is controlled, and F0 covaries,
4. Question: What's the mechanism responsible for F0 covariation?
5. Suggestions:
   a Cricothyroid contraction in [-voice] Cs to suppress voicing carries over into following Vs, raising their F0 (Löfqvist, et al., 1989; see also Halle & Stevens's, 1971, proposal for [stiff] versus [slack] folds),
   b Larynx lowering in [+voice] Cs to expand the oral cavity and slow the oral air pressure rise also lowers F0 (Hombert, et al., 1979; see Honda, et al. (1999): vocal folds are shortened by larynx lowering).

New data from English: Hanson (2009)

1. Voiced and voiceless stops and fricatives and nasal /m/ (baseline),
2. Voiceless stops: aspirated or unaspirated (in /s/-stop clusters),
3. High, low, and neutral F0 contexts (neutral = low),
4. 5 male and 5 female speakers.

New data from English: Hanson (2009)

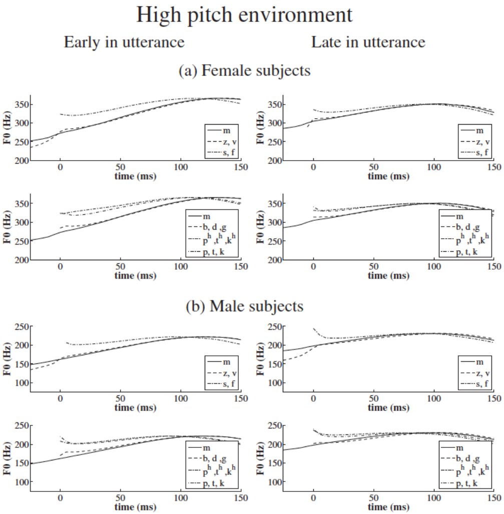
1. Voiced and voiceless stops and fricatives and nasal /m/ (baseline),
2. Voiceless stops: aspirated or unaspirated (in /s/-stop clusters),
3. High, low, and neutral F0 contexts (neutral = low),
4. 5 male and 5 female speakers.

New data from English: Hanson (2009)

1. Voiced and voiceless stops and fricatives and nasal /m/ (baseline),
2. Voiceless stops: aspirated or unaspirated (in /s/-stop clusters),
3. High, low, and neutral F0 contexts (neutral = low),
4. 5 male and 5 female speakers.

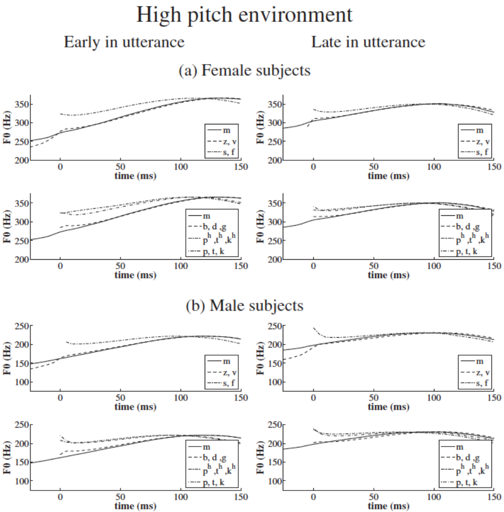
High context

1. s,f > z,v = m,
2. p,t,k,p^h,t^h,k^h > b,d,g = m

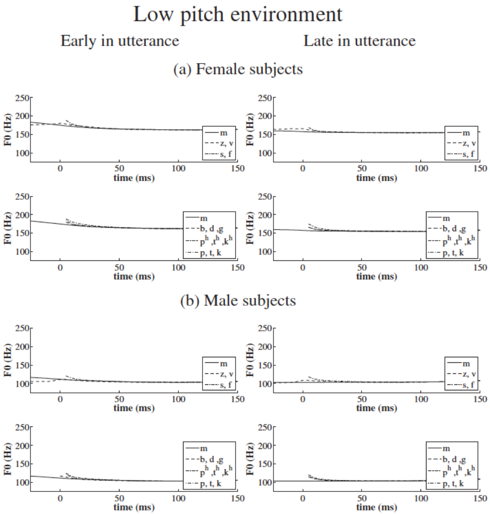


High context

1. $s,f > z,v = m,$
2. $p,t,k,p^h,t^h,k^h > b,d,g = m$

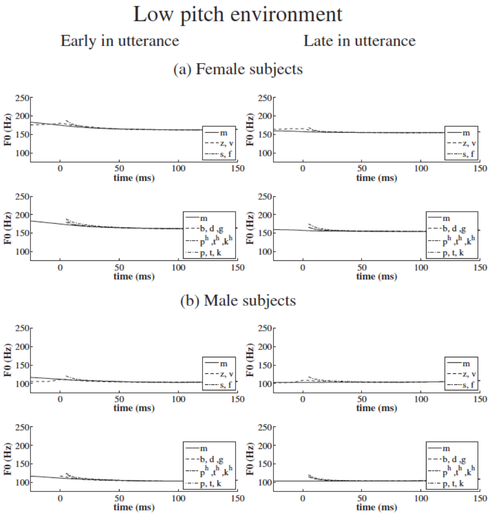


1. s,f $\geq$ z,v $=$ m,
2. p,t,k,p^h,t^h,k^h $\geq$ b,d,g $=$ m



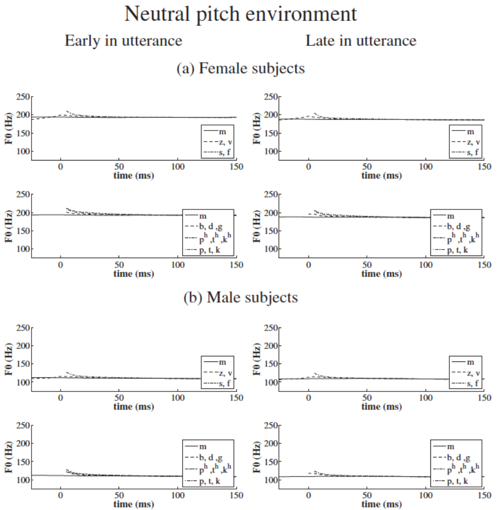
Low context

1. s,f $\geq$ z,v $=$ m,
2. p,t,k,p^h,t^h,k^h $\geq$ b,d,g $=$ m



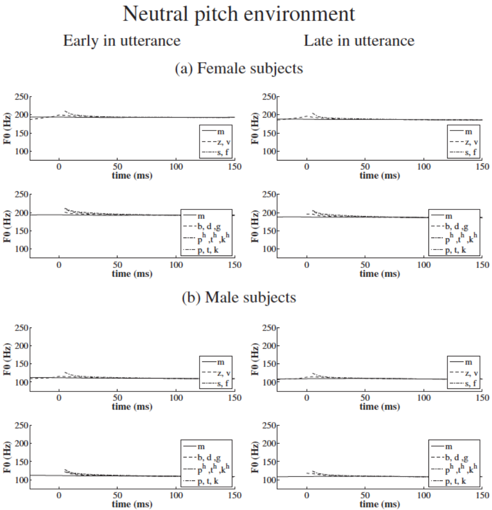
Neutral context

1. s,f $\geq$ z,v $=$ m,
2. p,t,k,p^h,t^h,k^h $\geq$ b,d,g $=$ m



Neutral context

1. s,f $\geq$ z,v $=$ m,
2. p,t,k,p^h,t^h,k^h $\geq$ b,d,g $=$ m



Hanson's interpretation

1. F0 is raised after voiceless obstruents unless doing so conflicts with producing a low F0 target,
2. Because F0 doesn't differ in all intonational contexts, it can't be produced by an independent gesture intended to enhance the [voice] contrast, contrary to Kingston & Diehl's (1994) claims.

Hanson's interpretation

1. F0 is raised after voiceless obstruents unless doing so conflicts with producing a low F0 target,
2. Because F0 doesn't differ in all intonational contexts, it can't be produced by an independent gesture intended to enhance the [voice] contrast, contrary to Kingston & Diehl's (1994) claims.

A response

1. Misinterpretation: Kingston & Diehl (1994) emphasize that the clearest evidence for articulatory control is systematic variation between contexts,
2. If F0 differences were a mechanical consequence of producing voicing versus voicelessness, then they would be observed across contexts,
3. 6/10 speakers produced equally high F0 contours after unaspirated stops following /s/ as aspirated stops, but 4/10 produced F0 contours intermediate between the higher aspirated and the lower unaspirated stops, /b,dg/, that don't follow /s/.
4. Kingston & Diehl (1994) reported similar findings (see also Ohde, 1984), and interpreted this variability as a side effect of neutralizing the [voice] contrast in this context.

A response

1. Misinterpretation: Kingston & Diehl (1994) emphasize that the clearest evidence for articulatory control is systematic variation between contexts,
2. If F0 differences were a mechanical consequence of producing voicing versus voicelessness, then they would be observed across contexts,
3. 6/10 speakers produced equally high F0 contours after unaspirated stops following /s/ as aspirated stops, but 4/10 produced F0 contours intermediate between the higher aspirated and the lower unaspirated stops, /b,dg/, that don't follow /s/.
4. Kingston & Diehl (1994) reported similar findings (see also Ohde, 1984), and interpreted this variability as a side effect of neutralizing the [voice] contrast in this context.

A response

1. Misinterpretation: Kingston & Diehl (1994) emphasize that the clearest evidence for articulatory control is systematic variation between contexts,
2. If F0 differences were a mechanical consequence of producing voicing versus voicelessness, then they would be observed across contexts,
3. 6/10 speakers produced equally high F0 contours after unaspirated stops following /s/ as aspirated stops, but 4/10 produced F0 contours intermediate between the higher aspirated and the lower unaspirated stops, /b,dg/, that don't follow /s/.
4. Kingston & Diehl (1994) reported similar findings (see also Ohde, 1984), and interpreted this variability as a side effect of neutralizing the [voice] contrast in this context.

A response

1. Misinterpretation: Kingston & Diehl (1994) emphasize that the clearest evidence for articulatory control is systematic variation between contexts,
2. If F0 differences were a mechanical consequence of producing voicing versus voicelessness, then they would be observed across contexts,
3. 6/10 speakers produced equally high F0 contours after unaspirated stops following /s/ as aspirated stops, but 4/10 produced F0 contours intermediate between the higher aspirated and the lower unaspirated stops, /b,dg/, that don't follow /s/.
4. Kingston & Diehl (1994) reported similar findings (see also Ohde, 1984), and interpreted this variability as a side effect of neutralizing the [voice] contrast in this context.

Hanson's interpretation and our response continued

1. The distinctive feature is [stiff],
2. When [+stiff] occurs in a low F0 context or [-stiff] in a high F0 context, these values are enhanced by vocal fold spreading and active vocal tract expansion, respectively,
3. But these enhancing gestures would produce voicelessness or voicing, not alter/enhance vocal fold stiffness itself,
4. Unexplained: F0 lowering after voiced obstruents would not be inhibited in low F0 contexts, yet F0 isn't lower after voiced obstruents than after /m/,
5. Acoustic correlates of the enhancing gestures aren't shown to differ between high and low F0 contexts (but see Hanson, 2004, for tentative evidence that voiced obstruents are less voiced in high F0 contexts).

Hanson's interpretation and our response continued

1. The distinctive feature is [stiff],
2. When [+stiff] occurs in a low F0 context or [-stiff] in a high F0 context, these values are enhanced by vocal fold spreading and active vocal tract expansion, respectively,
3. But these enhancing gestures would produce voicelessness or voicing, not alter/enhance vocal fold stiffness itself,
4. Unexplained: F0 lowering after voiced obstruents would not be inhibited in low F0 contexts, yet F0 isn't lower after voiced obstruents than after /m/,
5. Acoustic correlates of the enhancing gestures aren't shown to differ between high and low F0 contexts (but see Hanson, 2004, for tentative evidence that voiced obstruents are less voiced in high F0 contexts).

Hanson's interpretation and our response continued

1. The distinctive feature is [stiff],
2. When [+stiff] occurs in a low F0 context or [-stiff] in a high F0 context, these values are enhanced by vocal fold spreading and active vocal tract expansion, respectively,
3. But these enhancing gestures would produce voicelessness or voicing, not alter/enhance vocal fold stiffness itself,
4. Unexplained: F0 lowering after voiced obstruents would not be inhibited in low F0 contexts, yet F0 isn't lower after voiced obstruents than after /m/,
5. Acoustic correlates of the enhancing gestures aren't shown to differ between high and low F0 contexts (but see Hanson, 2004, for tentative evidence that voiced obstruents are less voiced in high F0 contexts).

Hanson's interpretation and our response continued

1. The distinctive feature is [stiff],
2. When [+stiff] occurs in a low F0 context or [-stiff] in a high F0 context, these values are enhanced by vocal fold spreading and active vocal tract expansion, respectively,
3. But these enhancing gestures would produce voicelessness or voicing, not alter/enhance vocal fold stiffness itself,
4. Unexplained: F0 lowering after voiced obstruents would not be inhibited in low F0 contexts, yet F0 isn't lower after voiced obstruents than after /m/,
5. Acoustic correlates of the enhancing gestures aren't shown to differ between high and low F0 contexts (but see Hanson, 2004, for tentative evidence that voiced obstruents are less voiced in high F0 contexts).

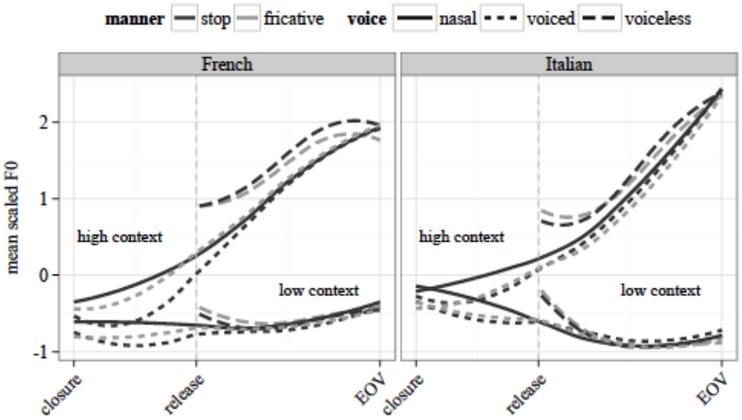
Hanson's interpretation and our response continued

1. The distinctive feature is [stiff],
2. When [+stiff] occurs in a low F0 context or [-stiff] in a high F0 context, these values are enhanced by vocal fold spreading and active vocal tract expansion, respectively,
3. But these enhancing gestures would produce voicelessness or voicing, not alter/enhance vocal fold stiffness itself,
4. Unexplained: F0 lowering after voiced obstruents would not be inhibited in low F0 contexts, yet F0 isn't lower after voiced obstruents than after /m/,
5. Acoustic correlates of the enhancing gestures aren't shown to differ between high and low F0 contexts (but see Hanson, 2004, for tentative evidence that voiced obstruents are less voiced in high F0 contexts).

Kirby & Ladd (ms.): Replicating Hanson (2009) in French and Italian

1. H: p,f $>$ m $=$ v $\geq$ b
2. L: p,f $>$ b,v,m
3. C H & L: m $>$ b,v

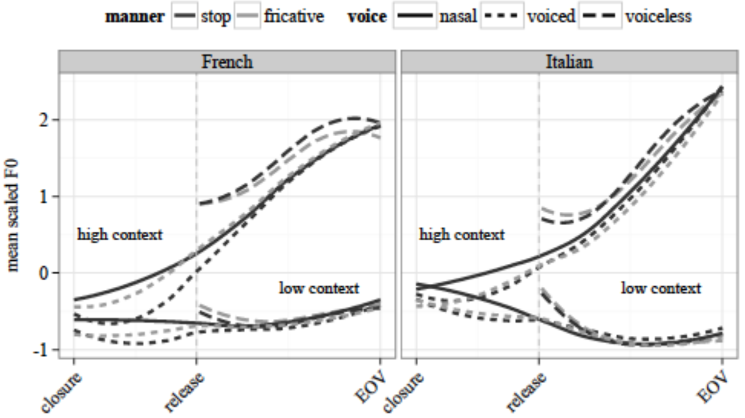
1. H: p,f $>$ m $\geq$ b,v
2. L: p,f $>$ b,v,m
3. C H & L: m $>$ b,v



Kirby & Ladd (ms.): Replicating Hanson (2009) in French and Italian

1. H: p,f $>$ m $=$ v $\geq$ b
2. L: p,f $>$ b,v,m
3. C H & L: m $>$ b,v

1. H: p,f $>$ m $\geq$ b,v
2. L: p,f $>$ b,v,m
3. C H & L: m $>$ b,v



Kirby & Ladd's interpretation

1. F0 is consistently raised following voiceless obstruents relative to after sonorants, by a vocal fold-stiffening gesture intended to inhibit voicing during the constriction, possibly CT contraction (Halle & Stevens, 1971; Löfqvist, 1989; see also Hanson, 2009),
2. F0 is consistently lowered during voiced obstruent constrictions relative to during sonorants, by an oral cavity expanding maneuver intended slow the build-up in oral air pressure (Honda, et al., 1999) – carries over into following Vs.

Kirby & Ladd's interpretation

1. F0 is consistently raised following voiceless obstruents relative to after sonorants, by a vocal fold-stiffening gesture intended to inhibit voicing during the constriction, possibly CT contraction (Halle & Stevens, 1971; Löfqvist, 1989; see also Hanson, 2009),
2. F0 is consistently lowered during voiced obstruent constrictions relative to during sonorants, by an oral cavity expanding maneuver intended slow the build-up in oral air pressure (Honda, et al., 1999) – carries over into following Vs.

A response

1. The [voice] contrast is realized differently in French and Italian than in English stops, as voicing lead versus short lag, rather than short lag versus long lag,
2. A voice-inhibiting adjustment such as vocal fold stiffening resulting from cricothyroid contraction would have to persist far longer in English than French or Italian to raise F0 substantially in the following vowel,
3. The much larger glottal opening for a English long lag stop would/should remove any need for such an inhibitory adjustment,
4. The [voice] contrast is realized similarly in French, Italian, and English fricatives, as voicing during the constriction versus its absence, yet the fricatives' perturb F0 similarly in all three languages to the stops.

A response

1. The [voice] contrast is realized differently in French and Italian than in English stops, as voicing lead versus short lag, rather than short lag versus long lag,
2. A voice-inhibiting adjustment such as vocal fold stiffening resulting from cricothyroid contraction would have to persist far longer in English than French or Italian to raise F0 substantially in the following vowel,
3. The much larger glottal opening for a English long lag stop would/should remove any need for such an inhibitory adjustment,
4. The [voice] contrast is realized similarly in French, Italian, and English fricatives, as voicing during the constriction versus its absence, yet the fricatives' perturb F0 similarly in all three languages to the stops.

A response

1. The [voice] contrast is realized differently in French and Italian than in English stops, as voicing lead versus short lag, rather than short lag versus long lag,
2. A voice-inhibiting adjustment such as vocal fold stiffening resulting from cricothyroid contraction would have to persist far longer in English than French or Italian to raise F0 substantially in the following vowel,
3. The much larger glottal opening for a English long lag stop would/should remove any need for such an inhibitory adjustment,
4. The [voice] contrast is realized similarly in French, Italian, and English fricatives, as voicing during the constriction versus its absence, yet the fricatives' perturb F0 similarly in all three languages to the stops.

A response

1. The [voice] contrast is realized differently in French and Italian than in English stops, as voicing lead versus short lag, rather than short lag versus long lag,
2. A voice-inhibiting adjustment such as vocal fold stiffening resulting from cricothyroid contraction would have to persist far longer in English than French or Italian to raise F0 substantially in the following vowel,
3. The much larger glottal opening for a English long lag stop would/should remove any need for such an inhibitory adjustment,
4. The [voice] contrast is realized similarly in French, Italian, and English fricatives, as voicing during the constriction versus its absence, yet the fricatives' perturb F0 similarly in all three languages to the stops.

Contour tones need more time

1. Contour tones may require two moras (Duanmu, 1990), or perhaps a syllable rime that lasts long enough (via phrase- or word-final lengthening) to realize the contour (Zhang, 2004),
2. In a balanced survery of 187 languages with contour tones (Zhang, 2002), 37 have only falling tones, while just 3 have only rising tones,
3. Rising tones often demand or create more sonorant material than falling tones, (Zhang, 2004), perhaps because it takes longer to produce a rise of a given size than a fall (Ohala & Ewan, 1973; Sundberg, 1973, 1979).

Contour tones need more time

1. Contour tones may require two moras (Duanmu, 1990), or perhaps a syllable rime that lasts long enough (via phrase- or word-final lengthening) to realize the contour (Zhang, 2004),
2. In a balanced survery of 187 languages with contour tones (Zhang, 2002), 37 have only falling tones, while just 3 have only rising tones,
3. Rising tones often demand or create more sonorant material than falling tones, (Zhang, 2004), perhaps because it takes longer to produce a rise of a given size than a fall (Ohala & Ewan, 1973; Sundberg, 1973, 1979).

Contour tones need more time

1. Contour tones may require two moras (Duanmu, 1990), or perhaps a syllable rime that lasts long enough (via phrase- or word-final lengthening) to realize the contour (Zhang, 2004),
2. In a balanced survery of 187 languages with contour tones (Zhang, 2002), 37 have only falling tones, while just 3 have only rising tones,
3. Rising tones often demand or create more sonorant material than falling tones, (Zhang, 2004), perhaps because it takes longer to produce a rise of a given size than a fall (Ohala & Ewan, 1973; Sundberg, 1973, 1979).

Intonation differs

1. Contour tones, especially rising ones, require more TBUs or longer stretches of tone-bearing material than level tones,
2. Prediction: When the tone-bearing material is shortened, F0 contours should be truncated not compressed = constant versus accelerated F0 change,
3. Confirmed for the majority of Swedish dialects (Alstermark & Erickson, 1971; Erikson & Alstermark, 1972; Bannert & Bredvad-Jensen, 1975),
4. Grabe (1998) shows that both falling and rising F0 contours are compressed in the Cambridge English, and that falling F0 contours are truncated in German, but rising ones are instead compressed,
5. Grabe, et al. (2000) shows that Newcastle English resembles Cambridge English in compressing both falling and rising F0 contours, while both are truncated in Leeds and Belfast English.

Intonation differs

1. Contour tones, especially rising ones, require more TBUs or longer stretches of tone-bearing material than level tones,
2. Prediction: When the tone-bearing material is shortened, F0 contours should be truncated not compressed = constant versus accelerated F0 change,
3. Confirmed for the majority of Swedish dialects (Alstermark & Erickson, 1971; Erikson & Alstermark, 1972; Bannert & Bredvad-Jensen, 1975),
4. Grabe (1998) shows that both falling and rising F0 contours are compressed in the Cambridge English, and that falling F0 contours are truncated in German, but rising ones are instead compressed,
5. Grabe, et al. (2000) shows that Newcastle English resembles Cambridge English in compressing both falling and rising F0 contours, while both are truncated in Leeds and Belfast English.

Intonation differs

1. Contour tones, especially rising ones, require more TBUs or longer stretches of tone-bearing material than level tones,
2. Prediction: When the tone-bearing material is shortened, F0 contours should be truncated not compressed = constant versus accelerated F0 change,
3. Confirmed for the majority of Swedish dialects (Alstermark & Erickson, 1971; Erikson & Alstermark, 1972; Bannert & Bredvad-Jensen, 1975),
4. Grabe (1998) shows that both falling and rising F0 contours are compressed in the Cambridge English, and that falling F0 contours are truncated in German, but rising ones are instead compressed,
5. Grabe, et al. (2000) shows that Newcastle English resembles Cambridge English in compressing both falling and rising F0 contours, while both are truncated in Leeds and Belfast English.

Intonation differs

1. Contour tones, especially rising ones, require more TBUs or longer stretches of tone-bearing material than level tones,
2. Prediction: When the tone-bearing material is shortened, F0 contours should be truncated not compressed = constant versus accelerated F0 change,
3. Confirmed for the majority of Swedish dialects (Alstermark & Erickson, 1971; Erikson & Alstermark, 1972; Bannert & Bredvad-Jensen, 1975),
4. Grabe (1998) shows that both falling and rising F0 contours are compressed in the Cambridge English, and that falling F0 contours are truncated in German, but rising ones are instead compressed,
5. Grabe, et al. (2000) shows that Newcastle English resembles Cambridge English in compressing both falling and rising F0 contours, while both are truncated in Leeds and Belfast English.

Intonation differs

1. Contour tones, especially rising ones, require more TBUs or longer stretches of tone-bearing material than level tones,
2. Prediction: When the tone-bearing material is shortened, F0 contours should be truncated not compressed = constant versus accelerated F0 change,
3. Confirmed for the majority of Swedish dialects (Alstermark & Erickson, 1971; Erikson & Alstermark, 1972; Bannert & Bredvad-Jensen, 1975),
4. Grabe (1998) shows that both falling and rising F0 contours are compressed in the Cambridge English, and that falling F0 contours are truncated in German, but rising ones are instead compressed,
5. Grabe, et al. (2000) shows that Newcastle English resembles Cambridge English in compressing both falling and rising F0 contours, while both are truncated in Leeds and Belfast English.

The puzzle

1. Why do the great majority of tone languages fail to compress contour tones, especially rising ones, when there are fewer TBUs or shorter stretches of tone-bearing material, while F0 contours, including rising ones, are compressed in languages where tones arise in the intonation rather than the lexicon?
2. Reframed: Why should speakers be able to control the speed of F0 change when the tones arise in the intonation but not when they arise in the lexicon?

The puzzle

1. Why do the great majority of tone languages fail to compress contour tones, especially rising ones, when there are fewer TBUs or shorter stretches of tone-bearing material, while F0 contours, including rising ones, are compressed in languages where tones arise in the intonation rather than the lexicon?
2. Reframed: Why should speakers be able to control the speed of F0 change when the tones arise in the intonation but not when they arise in the lexicon?

Audience design?

1. Do speakers adjust their articulations to make it easier for the listener to recognize what they've said?
2. Is articulatory control's purpose to design the phonetic contents of the speech signal to serve the audience?
3. Do combinations of controlled articulations and their acoustic consequences enhance minimal contrasts between speech sounds?
4. Possibly not: Cf. Lindblom's (1986) "sufficient dispersion" vs Liljencrantz & Lindblom's (1972) "maximal dispersion".

Audience design?

1. Do speakers adjust their articulations to make it easier for the listener to recognize what they've said?
2. Is articulatory control's purpose to design the phonetic contents of the speech signal to serve the audience?
3. Do combinations of controlled articulations and their acoustic consequences enhance minimal contrasts between speech sounds?
4. Possibly not: Cf. Lindblom's (1986) "sufficient dispersion" vs Liljencrantz & Lindblom's (1972) "maximal dispersion".

Audience design?

1. Do speakers adjust their articulations to make it easier for the listener to recognize what they've said?
2. Is articulatory control's purpose to design the phonetic contents of the speech signal to serve the audience?
3. Do combinations of controlled articulations and their acoustic consequences enhance minimal contrasts between speech sounds?
4. Possibly not: Cf. Lindblom's (1986) "sufficient dispersion" vs Liljencrantz & Lindblom's (1972) "maximal dispersion".

Audience design?

1. Do speakers adjust their articulations to make it easier for the listener to recognize what they've said?
2. Is articulatory control's purpose to design the phonetic contents of the speech signal to serve the audience?
3. Do combinations of controlled articulations and their acoustic consequences enhance minimal contrasts between speech sounds?
4. Possibly not: Cf. Lindblom's (1986) "sufficient dispersion" vs Liljencrantz & Lindblom's (1972) "maximal dispersion".

Audience design?

1. Bard et al. (2000) show that repeated mentions of a word in the Map Task are less clear than first mentions even when the speaker knew the listener hadn't heard the first mention, i.e. when new rather than given information,
2. Speakers control intelligibility at first and quickly only with reference to the information they themselves have, in which earlier mentions prime later ones,
3. Only later and more slowly do they draw inferences from a model of the listener's knowledge, often too slowly to alter the current articulatory plan (cf. Kahneman, 2011).

Audience design?

1. Bard et al. (2000) show that repeated mentions of a word in the Map Task are less clear than first mentions even when the speaker knew the listener hadn't heard the first mention, i.e. when new rather than given information,
2. Speakers control intelligibility at first and quickly only with reference to the information they themselves have, in which earlier mentions prime later ones,
3. Only later and more slowly do they draw inferences from a model of the listener's knowledge, often too slowly to alter the current articulatory plan (cf. Kahneman, 2011).

Audience design?

1. Bard et al. (2000) show that repeated mentions of a word in the Map Task are less clear than first mentions even when the speaker knew the listener hadn't heard the first mention, i.e. when new rather than given information,
2. Speakers control intelligibility at first and quickly only with reference to the information they themselves have, in which earlier mentions prime later ones,
3. Only later and more slowly do they draw inferences from a model of the listener's knowledge, often too slowly to alter the current articulatory plan (cf. Kahneman, 2011).

Coarticulation: A curse or a blessing?

1. Curse? Coarticulation produces enormous variability in the acoustic properties associated with distinctive feature values, and thus appears to make the listener's task harder,
2. Blessing? Coarticulation permits the speaker to produce a many speech sounds rapidly and fluently and thus convey considerable linguistic information in a short span of time,
3. How can the speaker get away with producing so much variability without making it impossible for the listener to recognize the distinctive feature values conveyed by the signal?
4. Apparently, because the listeners perceive invariance in/despite the variability,
5. How do listeners perceive invariance in variability?

Coarticulation: A curse or a blessing?

1. Curse? Coarticulation produces enormous variability in the acoustic properties associated with distinctive feature values, and thus appears to make the listener's task harder,
2. Blessing? Coarticulation permits the speaker to produce a many speech sounds rapidly and fluently and thus convey considerable linguistic information in a short span of time,
3. How can the speaker get away with producing so much variability without making it impossible for the listener to recognize the distinctive feature values conveyed by the signal?
4. Apparently, because the listeners perceive invariance in/despite the variability,
5. How do listeners perceive invariance in variability?

Coarticulation: A curse or a blessing?

1. Curse? Coarticulation produces enormous variability in the acoustic properties associated with distinctive feature values, and thus appears to make the listener's task harder,
2. Blessing? Coarticulation permits the speaker to produce a many speech sounds rapidly and fluently and thus convey considerable linguistic information in a short span of time,
3. How can the speaker get away with producing so much variability without making it impossible for the listener to recognize the distinctive feature values conveyed by the signal?
4. Apparently, because the listeners perceive invariance in/despite the variability,
5. How do listeners perceive invariance in variability?

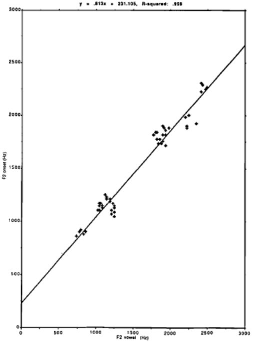
Coarticulation: A curse or a blessing?

1. Curse? Coarticulation produces enormous variability in the acoustic properties associated with distinctive feature values, and thus appears to make the listener's task harder,
2. Blessing? Coarticulation permits the speaker to produce a many speech sounds rapidly and fluently and thus convey considerable linguistic information in a short span of time,
3. How can the speaker get away with producing so much variability without making it impossible for the listener to recognize the distinctive feature values conveyed by the signal?
4. Apparently, because the listeners perceive invariance in/despite the variability,
5. How do listeners perceive invariance in variability?

Coarticulation: A curse or a blessing?

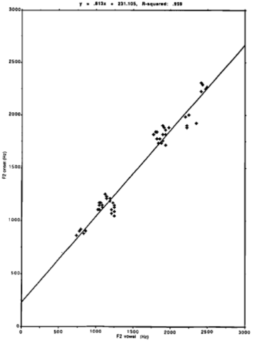
1. Curse? Coarticulation produces enormous variability in the acoustic properties associated with distinctive feature values, and thus appears to make the listener's task harder,
2. Blessing? Coarticulation permits the speaker to produce a many speech sounds rapidly and fluently and thus convey considerable linguistic information in a short span of time,
3. How can the speaker get away with producing so much variability without making it impossible for the listener to recognize the distinctive feature values conveyed by the signal?
4. Apparently, because the listeners perceive invariance in/despite the variability,
5. How do listeners perceive invariance in variability?

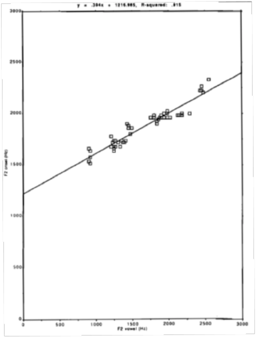
Bilabial = b
Steep slope
Low intercept
O = 0.813*V + 231



Locus equations (Sussman, et al., 1991): Lines fit to F2 Onset by F2 Vowel values

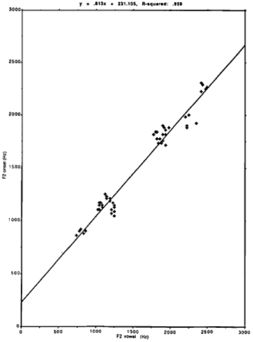
Bilabial = b	Alveolar = d
Steep slope	Shallow slope
Low intercept	High intercept
O = 0.813*V + 231	O = 0.394*V + 1217



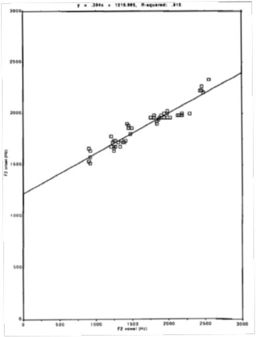


Locus equations (Sussman, et al., 1991): Lines fit to F2 Onset by F2 Vowel values

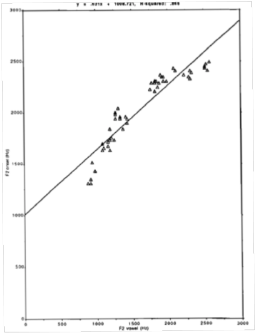
Bilabial = b
Steep slope
Low intercept
O = 0.813*V + 231



Alveolar = d
Shallow slope
High intercept
O = 0.394*V + 1217



Velar = g
Intermediate slope
Intermediate intercept
O = 0.631*V + 1009



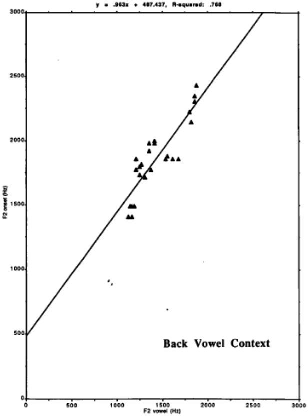
Locus equations:
Velars before back versus front vowels

Back Vowels = gv

Very steep slope

Low intercept

O = 0.963*V + 487



Articulatory Control?

Kingston

Jakobson, Fant, & Halle (1952)

Outline

Kingston & Diehl, 1994

New data

Time and F0 change

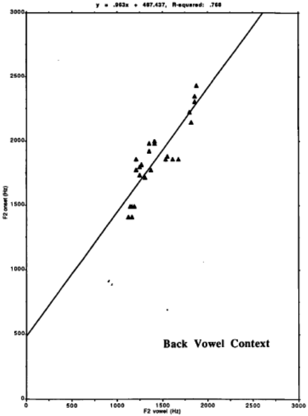
Audience design?

Coarticulation, sound change, and information distribution

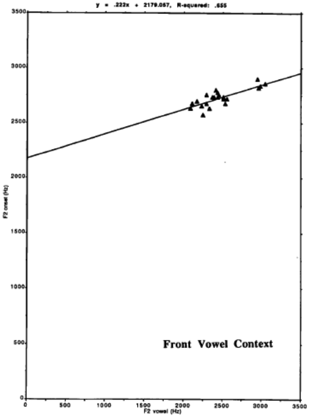
Summary

Locus equations: Velars before back versus front vowels

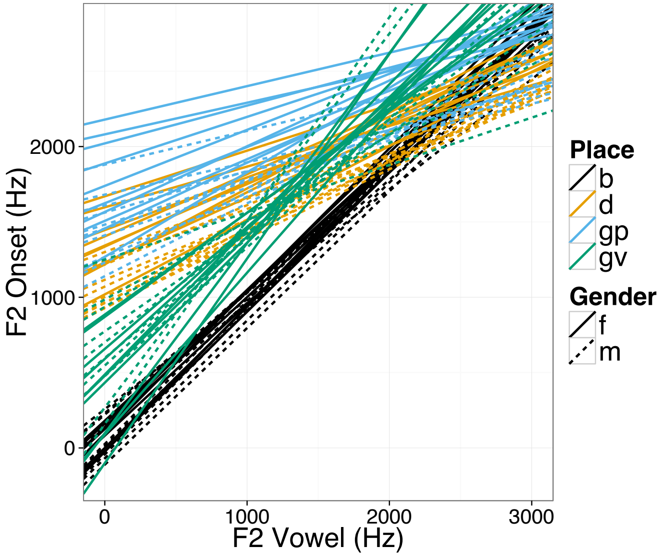
Back Vowels = gv
Very steep slope
Low intercept
O = 0.963*V + 487



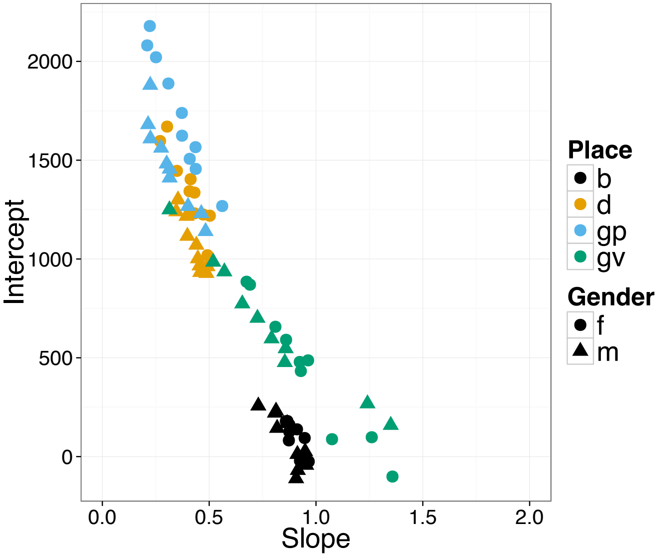
Front Vowels = gp
Very shallow slope
Very high intercept
O = 0.222*V + 2179



Locus equations: 10 women, 10 men (Sussman et al., 1991)



2nd order locus equations for 10 women, 10 men (Sussman et al., 1991)



Interim summary

1. First order:
   a. Slope = Coarticulation for backness with following Vs: b > g > d,
   b. Intercept = F2 onset minimum per place when F2 vowel = 0: b < g < d,
   c. Alternatively:
      i. Slope: gv > b > d > gp,
      ii. Intercept: b ≤ gv < d < gp;
2. Second order: Slope and intercept vary inversely.
   a. d, gp relatively invariant with vowel backness and high compared to b, gv,
   b. gp more variable than d, gv more variable than b.

Interim summary

1. First order:
   a. Slope = Coarticulation for backness with following Vs: b > g > d,
   b. Intercept = F2 onset minimum per place when F2 vowel = 0: b < g < d,
   c. Alternatively:
      i. Slope: gv > b > d > gp,
      ii. Intercept: b ≤ gv < d < gp;
2. Second order: Slope and intercept vary inversely.
   a. d, gp relatively invariant with vowel backness and high compared to b, gv,
   b. gp more variable than d, gv more variable than b.

Interim summary

1. First order:
   a. Slope = Coarticulation for backness with following Vs: b > g > d,
   b. Intercept = F2 onset minimum per place when F2 vowel = 0: b < g < d,
   c. Alternatively:
      i. Slope: gv > b > d > gp,
      ii. Intercept: b ≤ gv < d < gp;
2. Second order: Slope and intercept vary inversely.
   a. d, gp relatively invariant with vowel backness and high compared to b, gv,
   b. gp more variable than d, gv more variable than b.

Interim summary

1. First order:
   a. Slope = Coarticulation for backness with following Vs: $b > g > d$,
   b. Intercept = F2 onset minimum per place when F2 vowel = 0: $b < g < d$,
   c. Alternatively:
      i. Slope: $gv > b > d > gp$,
      ii. Intercept: $b \leq gv < d < gp$;
2. Second order: Slope and intercept vary inversely.
   a. d, gp relatively invariant with vowel backness and high compared to b, gv,
   b. gp more variable than d, gv more variable than b.

Frame-Content (Davis & MacNeilage, 1995)

1. Hypothesis: Babble is mandibular oscillation between closed ≈ consonantal and open ≈ vocalic states,
2. Prediction: V states homorganic with C states:
   a. Alveolar Cs–Front Vs,
   b. Velar–Back (unrounded) Vs,
   c. Labial–Central Vs (tongue at rest);
3. Transcription of 6 children's babble, from 6 to 12 months,
   a. Consonants:
      i. Labial (l): [p, b, m, w],
      ii. Alveolar (a): [t, d, n, j],
      iii. Velar (v): [k, g, ŋ];
   b. Vowels:
      i. Front (f): [i, ɪ, e, ɛ, æ],
      ii. Central (c): [a, ʌ, ə],
      iii. Back (b): [u, ʊ, o, ɔ]

Frame-Content (Davis & MacNeilage, 1995)

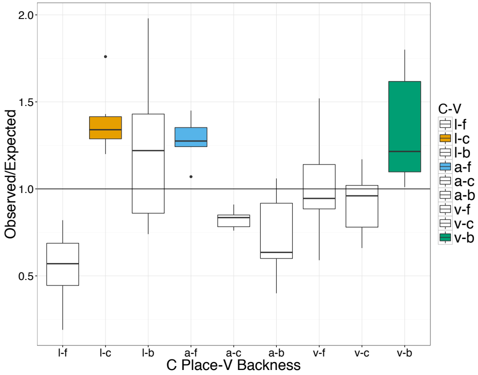
1. Hypothesis: Babble is mandibular oscillation between closed ≈ consonantal and open ≈ vocalic states,
2. Prediction: V states homorganic with C states:
   a. Alveolar Cs–Front Vs,
   b. Velar–Back (unrounded) Vs,
   c. Labial–Central Vs (tongue at rest);
3. Transcription of 6 children's babble, from 6 to 12 months,
   a. Consonants:
      i. Labial (l): [p, b, m, w],
      ii. Alveolar (a): [t, d, n, j],
      iii. Velar (v): [k, g, ŋ];
   b. Vowels:
      i. Front (f): [i, ɪ, e, ɛ, æ],
      ii. Central (c): [a, ʌ, ə],
      iii. Back (b): [u, ʊ, o, ɔ]

Frame-Content (Davis & MacNeilage, 1995)

1. Hypothesis: Babble is mandibular oscillation between closed ≈ consonantal and open ≈ vocalic states,
2. Prediction: V states homorganic with C states:
   a. Alveolar Cs–Front Vs,
   b. Velar–Back (unrounded) Vs,
   c. Labial–Central Vs (tongue at rest);
3. Transcription of 6 children's babble, from 6 to 12 months,
   a. Consonants:
      i. Labial (l): [p, b, m, w],
      ii. Alveolar (a): [t, d, n, j],
      iii. Velar (v): [k, g, ŋ];
   b. Vowels:
      i. Front (f): [i, ɪ, e, ɛ, æ],
      ii. Central (c): [a, ʌ, ə],
      iii. Back (b): [u, ʊ, o, ɔ]

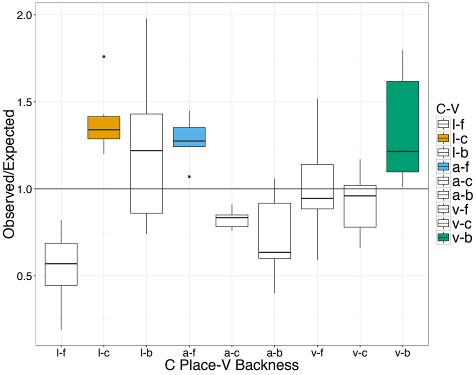
Observed/expected C-V cooccurrence ratios
Colors = O/E ratios predicted to exceed 1

1. Expected O/E > 1 for labial-central, alveolar-front, velar-back,
2. Labial-Back > 1,
3. Alveolar-Central, -Back < 1,
4. Velar-Front, -Central ≈ 1,
5. Vs coarticulate w/ Cs, not vice versa.



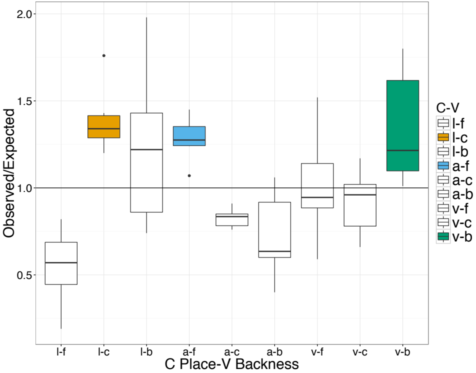
Observed/expected C-V cooccurrence ratios
Colors = O/E ratios predicted to exceed 1

1. Expected O/E > 1 for labial-central, alveolar-front, velar-back,
2. Labial-Back > 1,
3. Alveolar-Central, -Back < 1,
4. Velar-Front, -Central ≈ 1,
5. Vs coarticulate w/ Cs, not vice versa.

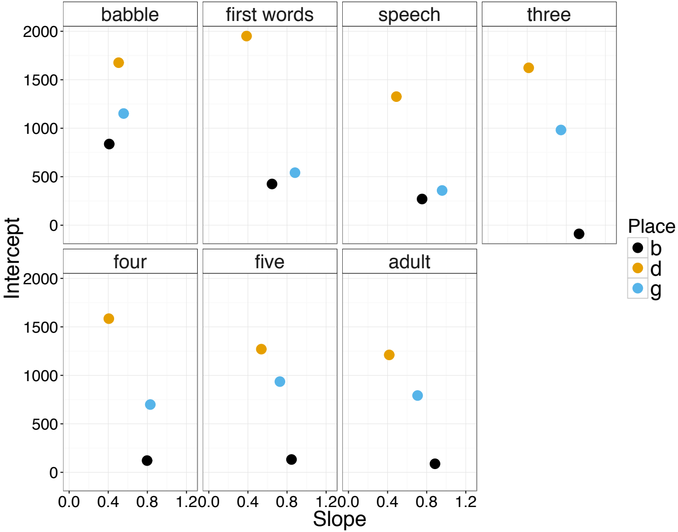


Observed/expected C-V cooccurrence ratios
Colors = O/E ratios predicted to exceed 1

1. Expected O/E > 1 for labial-central, alveolar-front, velar-back,
2. Labial-Back > 1,
3. Alveolar-Central, -Back < 1,
4. Velar-Front, -Central ≈ 1,
5. Vs coarticulate w/ Cs, not vice versa.

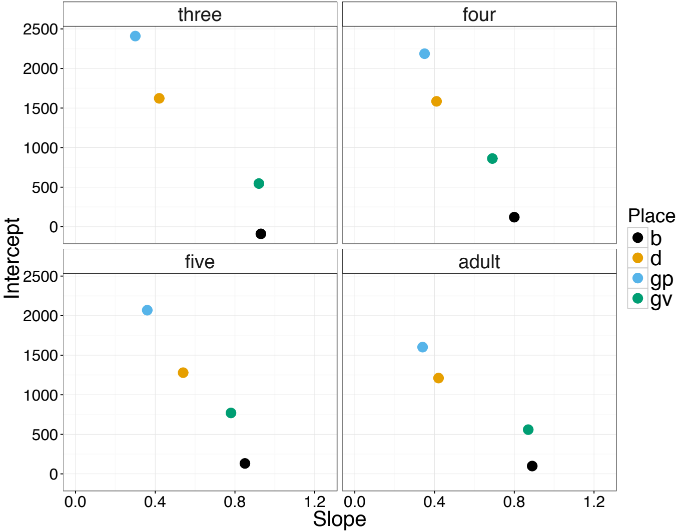


Acquiring locus equations: Babble to adulthood
Second order: b, d, g



Articulatory Control?
Kingston
Jakobson, Fant, & Halle (1952)
Outline
Kingston & Diehl, 1994
New data
Time and F0 change
Audience design?
Coarticulation, sound change, and information distribution
Summary

Acquiring locus equations: 3 years to adulthood
Second order: b, d, gp, gv



Acquisition summary

1. Before the age of 3:
   a Children's stops partially resemble adults' stops in F2 onset, intercepts: b $<$ d,
   b Children's stops in first words and speech resemble adults' in the extent of coarticulation with vowels' backness, slopes: b $>$ d,
   c Children's g doesn't resemble adults' in intercept or slope;
2. From the age of 3, slopes and intercepts for all places, including the difference between gp and gv is adult-like,
3. Iskarous, et al. (2010). Locus equations represent differences between consonant places in the extent to which their production permits coarticulation with tongue backness in vowels: gp $<$ d $\ll$ gv $<$ b,
4. Place-specific control of C-to-V coarticulation is learned.

Acquisition summary

1. Before the age of 3:
   a. Children's stops partially resemble adults' stops in F2 onset, intercepts: b $<$ d,
   b. Children's stops in first words and speech resemble adults' in the extent of coarticulation with vowels' backness, slopes: b $>$ d,
   c. Children's g doesn't resemble adults' in intercept or slope;
2. From the age of 3, slopes and intercepts for all places, including the difference between gp and gv is adult-like,
3. Iskarous, et al. (2010). Locus equations represent differences between consonant places in the extent to which their production permits coarticulation with tongue backness in vowels: gp $<$ d $\ll$ gv $<$ b,
4. Place-specific control of C-to-V coarticulation is learned.

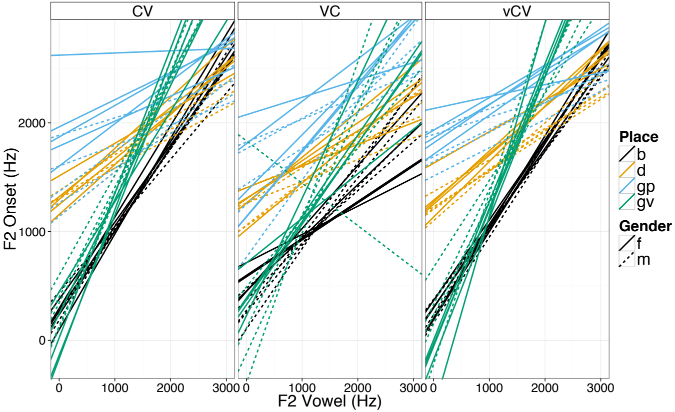
Acquisition summary

1. Before the age of 3:
   a. Children's stops partially resemble adults' stops in F2 onset, intercepts: b $<$ d,
   b. Children's stops in first words and speech resemble adults' in the extent of coarticulation with vowels' backness, slopes: b $>$ d,
   c. Children's g doesn't resemble adults' in intercept or slope;
2. From the age of 3, slopes and intercepts for all places, including the difference between gp and gv is adult-like,
3. Iskarous, et al. (2010). Locus equations represent differences between consonant places in the extent to which their production permits coarticulation with tongue backness in vowels: gp $<$ d $\ll$ gv $<$ b,
4. Place-specific control of C-to-V coarticulation is learned.

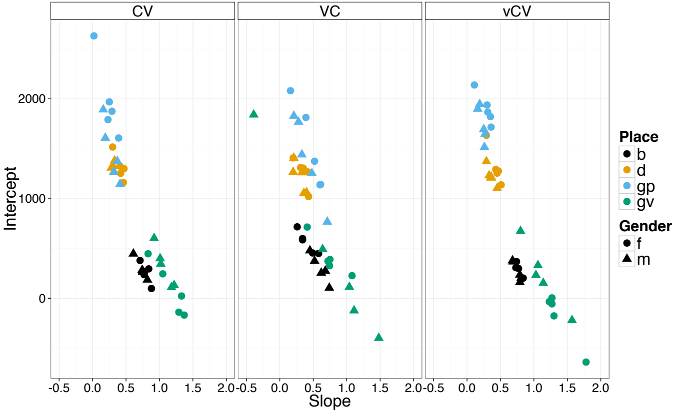
Acquisition summary

1. Before the age of 3:
   a. Children's stops partially resemble adults' stops in F2 onset, intercepts: b $<$ d,
   b. Children's stops in first words and speech resemble adults' in the extent of coarticulation with vowels' backness, slopes: b $>$ d,
   c. Children's g doesn't resemble adults' in intercept or slope;
2. From the age of 3, slopes and intercepts for all places, including the difference between gp and gv is adult-like,
3. Iskarous, et al. (2010). Locus equations represent differences between consonant places in the extent to which their production permits coarticulation with tongue backness in vowels: gp $<$ d $\ll$ gv $<$ b,
4. Place-specific control of C-to-V coarticulation is learned.

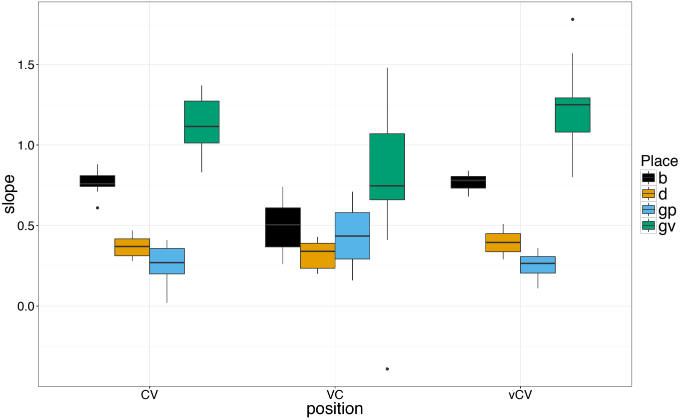
Locus equations by syllable position: 5 women, 5 men (Sussman et al., 1997)



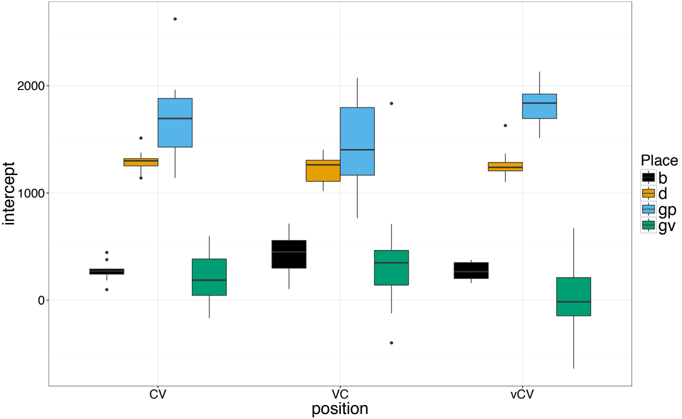
2nd order locus equations by syllable position: 5 women, 5 men (Sussman et al., 1997)



Variation by position: Slope



Variation by position: Intercept



VC offset versus CV and vCV onset

1. Slope:
   a VC $<$ CV, vCV, except gp,
   b VC more variable than CV, vCV;
2. Intercept:
   a Places differ less for VC than CV, vCV,
   b VC more variable than CV, vCV, except gv;

VC offset versus CV and vCV onset

1. Slope:
   a VC $<$ CV, vCV, except gp,
   b VC more variable than CV, vCV;
2. Intercept:
   a Places differ less for VC than CV, vCV,
   b VC more variable than CV, vCV, except gv;

Mutual information (Iskarous, et al., 2013)

1. The extent to which one articulation coarticulates with = depends on another, it conveys the same information as the other,
2. Locus equations measure both the dependence of the acoustics of consonant place on vowel backness, in their slopes, and the independence of place acoustics, in their intercepts:
   a. Dependence: b, gv $>$ d, gp,
   b. Independence: b, gv $\neq$ d, gp;
3. Both dependence and independence are weaker in codas (VC) than onsets (CV, vCV).

Mutual information (Iskarous, et al., 2013)

1. The extent to which one articulation coarticulates with = depends on another, it conveys the same information as the other,
2. Locus equations measure both the dependence of the acoustics of consonant place on vowel backness, in their slopes, and the independence of place acoustics, in their intercepts:
   a. Dependence: b, gv $>$ d, gp,
   b. Independence: b, gv $\neq$ d, gp;
3. Both dependence and independence are weaker in codas (VC) than onsets (CV, vCV).

Mutual information (Iskarous, et al., 2013)

1. The extent to which one articulation coarticulates with = depends on another, it conveys the same information as the other,
2. Locus equations measure both the dependence of the acoustics of consonant place on vowel backness, in their slopes, and the independence of place acoustics, in their intercepts:
   a Dependence: b, gv $>$ d, gp,
   b Independence: b, gv $\neq$ d, gp;
3. Both dependence and independence are weaker in codas (VC) than onsets (CV, vCV).

(In)dependence of information and sound change

1. Listeners misparse VC transitions as information about vowel backness rather than consonant place (Kingston, et al., in prep.),
2. Misparsing occurs when acoustic properties of a later sound are treated as information about the current one,
3. Misparsing is more likely to lead to sound change when acoustic dependence is greater. Place assimilation (Bybee & Easterday, in prep):
   a. C-to-V: Anticipatory (CV) 86 versus perseverative (VC) 17,
   b. V-to-C: Anticipatory (VC) 39 versus perseverative (CV) 25.
4. Larger proportion of anticipatory assimilation when coarticulation is greater, in CV than VC.

(In)dependence of information and sound change

1. Listeners misparse VC transitions as information about vowel backness rather than consonant place (Kingston, et al., in prep.),
2. Misparsing occurs when acoustic properties of a later sound are treated as information about the current one,
3. Misparsing is more likely to lead to sound change when acoustic dependence is greater. Place assimilation (Bybee & Easterday, in prep):
   a. C-to-V: Anticipatory (CV) 86 versus perseverative (VC) 17,
   b. V-to-C: Anticipatory (VC) 39 versus perseverative (CV) 25.
4. Larger proportion of anticipatory assimilation when coarticulation is greater, in CV than VC.

(In)dependence of information and sound change

1. Listeners misparse VC transitions as information about vowel backness rather than consonant place (Kingston, et al., in prep.),
2. Misparsing occurs when acoustic properties of a later sound are treated as information about the current one,
3. Misparsing is more likely to lead to sound change when acoustic dependence is greater. Place assimilation (Bybee & Easterday, in prep):
   a. C-to-V: Anticipatory (CV) 86 versus perseverative (VC) 17,
   b. V-to-C: Anticipatory (VC) 39 versus perseverative (CV) 25.
4. Larger proportion of anticipatory assimilation when coarticulation is greater, in CV than VC.

Summary

1. F0 is controlled following obstruents contrasting for [voice],
2. The timing of F0 change is controlled differently in intonation than tone,
3. Speakers don't design production to serve their audience's need for information,
4. Control of coarticulation must be learned, but even once learned can lead to sound change when information (about place) is shared between successive sounds.

Summary

1. F0 is controlled following obstruents contrasting for [voice],
2. The timing of F0 change is controlled differently in intonation than tone,
3. Speakers don't design production to serve their audience's need for information,
4. Control of coarticulation must be learned, but even once learned can lead to sound change when information (about place) is shared between successive sounds.

Summary

1. F0 is controlled following obstruents contrasting for [voice],
2. The timing of F0 change is controlled differently in intonation than tone,
3. Speakers don't design production to serve their audience's need for information,
4. Control of coarticulation must be learned, but even once learned can lead to sound change when information (about place) is shared between successive sounds.

Summary

1. F0 is controlled following obstruents contrasting for [voice],
2. The timing of F0 change is controlled differently in intonation than tone,
3. Speakers don't design production to serve their audience's need for information,
4. Control of coarticulation must be learned, but even once learned can lead to sound change when information (about place) is shared between successive sounds.