Assimilation from the listener’s perspective: adult and toddler data

Sharon Peperkamp

Laboratoire de Sciences Cognitives et Psycholinguistique
Institut d’Etude de la Cognition
Ecole Normale Supérieure
Paris
Introduction

• Assimilation, i.e. the spreading of one or several phonetic features from one speech sound to an adjacent one, is cross-linguistically very common.

• Two examples:
  – English:
    • regressive place assimilation of /t,d,n/ before labials and velars
    • sweet boy $\rightarrow$ swee[p] boy
  – French:
    • regressive voice assimilation in obstruent clusters
    • mec doux $\rightarrow$ me[g] doux ‘soft guy’
Introduction

- Listeners have detailed knowledge of assimilation and use it to ‘undo’ its effect during word recognition.
  - **cross-modal priming**: the written word *LEAN* is primed by an auditory sentence containing ‘lea[m] bacon’ but not by one containing ‘lea[m] game’ (Marslen-Wilson et al., 1995, Snoeren et al., 2008).
  - **word detection**: ‘sweet’ is recognized in an auditorily presented sentence containing ‘swee[p] boy’ but not in one containing ‘swee[p] dog’; this holds even when assimilation is complete (Darcy et al., 2009).

- To a large extent, compensation is language-specific
  - listeners show no or less compensation for non-native assimilation (Mitterer et al., 2006; Darcy et al., 2007).
Research questions

• What is the time course of compensation?
  – does compensation reflect a late and explicit correction of assimilated sounds that are initially categorized as such?
  – or does it occur at an early, automatic, perceptual stage, such that the assimilated sound is directly categorized as the intended sound?

• When and how is assimilation acquired by young children?
  – do children show adult-like compensation early on?
  – what type of learning algorithm do they exploit?
Part 1:
The time course of compensation for assimilation in adults
Case study: French voice assimilation

• Most assimilation rules are completely straightforward from a phonetic point of view
  – spreading of a phonetic feature if and only if it is present
    • swee[p] boy: leftward spreading of labial
    • * swee[p] dog: no labial, hence no spreading

• French voice assimilation is different
  – leftward spreading from obstruents but not from sonorants
    • clu[b] [s]uisse → clu[p] [s]uisse ‘Swiss club’
    • bu[s] [v]ert → bu[z] [v]ert ‘green bus’
    • bu[s] [n]oir → * bu[z] [n]oir ‘black bus’
  – compensation for the [-voice] → [+voice] change must involve a phonological (not just a phonetic) mechanism
Case study: French voice assimilation

• Discrimination task
  – due to compensation, listeners might have difficulty discriminating between assimilated and non-assimilated sequences
    e.g., ‘bu[s] vert’ vs. ‘bu[z] vert’: perceived as identical
  – however, they should have no difficulty discriminating sequences in which the feature change cannot be due to assimilation
    e.g., ‘bu[s] noir’ vs. ‘bu[z] noir’: perceived as different
Case study: French voice assimilation

• Use EEG (electroencephalography)
  – high temporal resolution (~ms)
  – relevant component: Mismatch Negativity (MMN)
    • an electrophysiological marker of automatic detection of deviance in a series of stimuli
    • arises very early, from around 150ms after the onset of deviance
    • sensitive to phonological (and not just acoustic) changes (Näätänen et al. 1997; Dehaene-Lambertz, 1997)
Case study: French voice assimilation

• Predictions

  – if compensation reflects a late and explicit correction of assimilated sounds that are initially categorized as such, then we should see the presence of an MMN for a feature change, regardless of whether it is induced by assimilation

  – if compensation occurs at an early perceptual stage, i.e. if the assimilated sound is directly categorized as the intended sound, then we should see

    • the absence of an MMN for a feature change that is induced by assimilation

    • presence of an MMN for a feature change that is not induced by assimilation
Design

Participants: French monolinguals

Task: Mismatch detection

<table>
<thead>
<tr>
<th>Precursor stimuli</th>
<th>Test stimulus</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>A</td>
<td>B</td>
</tr>
</tbody>
</table>

4 female speakers

1 male speaker

EEG recording:

high-density system (256 electrodes)
Stimuli

- 4 quadruplets of the form $V_1C_1C_2V_2$:

<table>
<thead>
<tr>
<th>$C_1$ voiceless obstruent</th>
<th>[usda]</th>
<th>[ɔfbe]</th>
<th>[ikdo]</th>
<th>[akvi]</th>
<th>$C_2$ voiced obstruent</th>
</tr>
</thead>
<tbody>
<tr>
<td>[uzda]</td>
<td>[ɔvbe]</td>
<td>[igdo]</td>
<td>[agvi]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_1$ voiced obstruent</td>
<td>[usra]</td>
<td>[ɔfne]</td>
<td>[ikmo]</td>
<td>[akni]</td>
<td>$C_2$ sonorant</td>
</tr>
<tr>
<td>[uzra]</td>
<td>[ɔvne]</td>
<td>[igmo]</td>
<td>[agni]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- multiple tokens by 1 male and 4 female speakers

- cross-splicing to create stimuli

- voiced obstruents are always produced as such by speakers (hence, complete assimilation)
# Predictions

<table>
<thead>
<tr>
<th>Context for assimilation</th>
<th>Condition</th>
<th>Example</th>
<th>Expected behavioral response</th>
<th>Expected physiological response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viable</td>
<td>control</td>
<td>ak vi – ak vi</td>
<td>same</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ag vi – ag vi</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>deviant</td>
<td>ak vi – ag vi</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ag vi – ak vi</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>Unviable</td>
<td>control</td>
<td>ak ni – ak ni</td>
<td>same</td>
<td>MMN</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ag ni – ag ni</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>deviant</td>
<td>ak ni – ag ni</td>
<td>different</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ag ni – ak ni</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Behavioral results

Error rate

- Viable: 
  - No-change
  - Voice-change

- Unviable: 
  - No-change
  - Voice-change

Context X Condition: $p < 0.001$

Reaction time

- Viable: 
  - No-change
  - Voice-change

- Unviable: 
  - No-change
  - Voice-change

Context X Condition: $p < 0.01$

Sun et al., in prep.
Electrophysiological results

1st response: 130 – 240 ms

Unviable

Viable

Fronto-central negativity:

Mismatch negativity

Context X Condition: $p < 0.05$

Sun et al., in prep.
Electrophysiological results

2\textsuperscript{nd} response: 360 – 520 ms

Unviable

Viable

Parietal positivity:

P300

Context X Condition: \( p < 0.001 \)

Sun et al., in prep.
Electrophysiological results

Significant ERP clusters over time
(voice-change minus no-change)

Sun et al., in prep.
Discussion

• Behavioral results:
  – listeners have difficulty perceiving a voicing change, but only when it is licensed by assimilation

• EEG results:
  – two responses related to the detection of voicing change in the unviable context only
    • MMN: automatic detection of voice deviation
    • P300: decision making process
Discussion

• What exactly did listeners perceive in the viable assimilation condition?
  – perceptual compensation of assimilation:
    [agvi] perceived as [akvi]
  – perceptual repair of phonotactic constraint violation, as within words, consonant clusters have to agree in voicing:
    [akvi] perceived as [agvi] or as [akfi]

• Transcription task
  – all precursor and test stimuli
  – coding of errors:
    • compensation: [+vce] [+vce] transcribed as [-vce] [+vce]
    • repair: [-vce] [+vce] transcribed as [+vce] [+vce] or as [-vce] [-vce]
Discussion

[agvi] perceived as [akvi]; listeners compensate for assimilation.
Discussion

• The absence of the MMN in viable context shows that compensation relies on early perceptual mechanisms (see also Mitterer et al., 2003, 2006; Tavabi et al., 2009)

• These mechanisms cannot be reduced to compensation for co-articulation, as both the viable and the unviable contexts involve voiced segments (i.e. voiced obstruents vs. sonorants)

→ evidence for an early role of the phonological grammar
Part 2: toddlers
Introduction

• Previous research on the acquisition of assimilation:

• Production:
  – case study of a British English child, producing place assimilation at age 2;10 (Newton and Wells, 2002)

• Perception:
  – Dutch 8-year-olds and English 7-year-olds compensate for native place assimilation (Blomert et al., 2004; Marshall et al., 2010)
Introduction

• At what age do children show adult-like compensation?
  – compensation for native assimilation
  – no (or less) compensation for non-native assimilation

• Test toddlers’ processing of both native and non-native assimilation
  – **participants**: 33-month-old French toddlers
  – **assimilation**: voice (native) and place (non-native)
  – **task**: forced-choice picture-pointing
“Voici un bus.”
“Et voilà un buz.”
control_familiar  Montre le bu[s]  Show the bus
control_unfamiliar  Montre le bu[z]  Show the bu[z]

viable assimilation  Montre le bu[z] de Paul  Show Paul’s bus/[z]

unviable assimilation  Montre le bu[z] là-bas  Show the bu[z] over there
Stimuli

- 12 obstruent-final French nouns known by 33-month-olds.
- Recordings by a female French speaker in child-directed speech.
- Speaker was instructed not to produce any pauses between words and to produce complete voicing changes.
- Evaluation of critical consonants by adult native listeners:
  - **stimuli**: final VC portions of target words
  - **task**: 2-alternative forced choice
  - **results**:
    - viable: 92.3% response assimilated consonant
    - unviable: 96.5% response assimilated consonant
Skoruppa, Mani & Peperkamp (2013)
“Ceci est une lune.”
“Et ça, c’est une lume.”
| control_familiar  | Montre la [lyn]           | Show the moon                      |
| control_unfamiliar| Montre la [lym]           | Show the [lym]                      |
| viable assimilation | Montre la [lym] par ici   | Show the moon over here            |
| unviable assimilation | Montre la [lym] de Paul  | Show Paul’s moon/[lym]             |
Stimuli

• 12 French nouns ending in a coronal obstruent or nasal, known by 33-month-old.

• Same speaker as in previous experiment.

• Evaluation of critical consonants by adult native listeners:
  – **stimuli**: final VC portions of target words
  – **task**: 2-alternative forced choice
  – **results**:
    • viable: 98.6% response assimilated consonant
    • unviable: 96.5% response assimilated consonant
% choice of familiar object

0 20 40 60 80 100

Control

Experimental

Skoruppa, Mani & Peperkamp (2013)
voice change

<table>
<thead>
<tr>
<th>Voice Change</th>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiar</td>
<td><img src="chart1.png" alt="Chart" /></td>
<td><img src="chart2.png" alt="Chart" /></td>
</tr>
<tr>
<td>Unfamiliar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viable</td>
<td><img src="chart3.png" alt="Chart" /></td>
<td><img src="chart4.png" alt="Chart" /></td>
</tr>
<tr>
<td>Unviable</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

place change

<table>
<thead>
<tr>
<th>Place Change</th>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiar</td>
<td><img src="chart5.png" alt="Chart" /></td>
<td><img src="chart6.png" alt="Chart" /></td>
</tr>
<tr>
<td>Unfamiliar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viable</td>
<td><img src="chart7.png" alt="Chart" /></td>
<td><img src="chart8.png" alt="Chart" /></td>
</tr>
<tr>
<td>Unviable</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Discussion

• French 33-month-olds compensate for native voice assimilation but not for non-native place assimilation.

• Would younger toddlers show the same behavior?
  – 24-month-olds have smaller lexicons
  – they are sensitive to word-final phonetic changes, but so far this has only been tested utterance-finally

“Where’s the boat?”
“Where’s the boad?”

Less fixations of the target (boat) in response to a mispronunciation (boa[d]) (Swingley, 2009)
Next experiments

• Test 24-month-olds’ processing of both native and non-native assimilation
  – **participants**: French and English
  – **assimilation**: voice (native for French, non-native for English)
  – **paradigm**: preferential looking
“Une tasse !”
A cup!
“Une ta[z] !”
control: Regarde la ta[s] maintenant. Look at the cup now.

viable: Regarde la ta[z] devant toi. Look at the cup in front of you.

unviable: Regarde la ta[z] maintenant. Look at the ta[z] now.
Stimuli

- 15 obstruent-final French nouns known by 24-month-olds.
- Recordings by a balanced French-English bilingual in child-directed speech.
- Evaluation of critical consonants by adult native listeners in 2AFC task on final VC portions of target words:

```
"tasse"
ta/s/
```
Skoruppa et al., in press, *Infancy*
“Here’s a bus!”
“And there’s a bu[z]!”
control: Can you find the bu[s] now?
visible: Can you find the bu[z] there?
unviable: Can you find the bu[z] now?
Stimuli

- 15 obstruent-final English nouns known by 24-month-olds.

- Recordings in child-directed speech by the same balanced French-English bilingual as in the previous experiment.

- Evaluation of critical consonants by adult native listeners in 2AFC task on final VC portions of target words:

```
bus
```

```
[s]
```

```
[z]
```

```
now
```

```
there
```

```
now
```

**Mean % of choice base consonant**

- **control**
- **viable**
- **unviable**
Mean proportion of looking to familiar object

- Standard
- Viable
- Unviable

Skoruppa et al., in press, *Infancy*
## Overview of results

<table>
<thead>
<tr>
<th>33 months</th>
<th>native</th>
<th>non-native</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>voice</td>
<td>place *</td>
</tr>
<tr>
<td>English</td>
<td>place</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>24 months</th>
<th>native</th>
<th>non-native</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>voice</td>
<td></td>
</tr>
<tr>
<td>French</td>
<td>voice</td>
<td></td>
</tr>
</tbody>
</table>
Discussion

• Toddlers compensate for native but not for non-native assimilation.

• What is the initial state, and hence what exactly have they acquired?
  – ‘empty’ grammar: no assimilation
    → native assimilation has been learned
  – ‘phonetically natural’ grammar: all sorts of assimilation, native and non-native (Donegan & Stampe, 1978)
    → non-native assimilations have been unlearned
Discussion

• Current study:
  – EEG with French 14-month-olds

• Preliminary results:
  – MMN in both the viable and the unviable context, hence infants detect the voicing change regardless of context

• If confirmed, this means:
  – no assimilation at the initial state
  – native assimilation is acquired between 14 and 24 months of age
Discussion

• How is assimilation acquired?

• Let’s first look at the input: rate of assimilation in child-directed speech
  – French Childes corpora: 55 000 short utterances containing approximately 216 000 words
  – around 1% of words are assimilable
    • end in an obstruent
    • are followed by a word starting with an obstruent with opposite voicing value
    • are not separated from the following word by a pause
  – note: unknown how many of these have actually undergone voice assimilation
Discussion

• How is assimilation acquired?

• One possibility: a lexical learning algorithm
  – knowing the meaning of words allows toddlers to observe the change in final consonant in certain contexts
    • bu[s] noir ‘black bus’ – bu[z] vert ‘green bus’

• How likely is this algorithm?
  – toddlers would need to know a ‘sufficient’ number of obstruent-final words and encounter them ‘sufficiently’ often in both their canonical and their assimilated forms
Discussion

• Occurrence of words that toddlers know?
  – 12 test items used with French 24-month-olds
  – 738 tokens in the Childes corpus (0.3% of entire corpus)
→ less than 0.1 % of word tokens in the input are known by 24-month-olds and might have undergone assimilation
• Multiple tokens of lexical items tend to occur in close distance to one another, and this so-called ‘burstiness’ helps lexical acquisition (Altmann et al., 2009).

• Burstiness could also help the acquisition of assimilation, as it makes canonical and assimilated tokens cluster together:

Ex.: *singe /sɛʒ/ ‘monkey’*  
*C'est le singe [sɛʒ] qui veut monter.*  
‘It’s the monkey who wants to go up.

*Regarde, cherche le singe [sɛʒ].*  
Look, find the monkey.

*Le singe [sɛʒ] est dans l'arbre, sur les bras.*  
The monkey is in the tree, on its arms.’

*Les petits singes [sɛʒ], ce sont des singes [sɛʃ] qui font les petits singes [sɛʒ].* ‘(As to) baby monkeys, it’s monkeys who make baby monkeys.’
Discussion

• An alternative prelexical learning algorithm
  – canonical and assimilated word forms have complementary distributions
    • $bu[s]$ occurs before voiceless obstruents
    • $bu[z]$ occurs before voiced obstruents, sonorants, vowels, and in final position.
  – final-voicing minimal pairs are problematic for this account
    • $bac$ (ba$[k]$) ‘ferry’ – $bague$ (ba$[g]$) ‘ring’
    • but such pairs are rare in the toddler’s input
  – prerequisites for this algorithm:
    • segmenting and storing word forms
      → 11-month-old infants have started doing this (Hallé & De Boisson-Bardies, 1994, 1996; Vihman et al., 2004; Swingley, 2005)
    • tracking of complementary distributions
      → 12-month-old infants can do this (White et al., 2008)
Conclusion

• Compensation for native assimilation takes place at an early processing stage, even if it cannot be reduced to compensation for co-articulation.

• Thus, the phonological grammar modulates early auditory brain responses.

• Toddlers as young as 24 months show adult-like processing of assimilation.
  – They compensate for native assimilation.
  – They do not compensate for non-native assimilation.

• Exactly when and how they acquire assimilation remain open questions.
Acknowledgments

**Adults:** Yue Sun

Martine Adda-Decker, Leonardo Barbosa, Maria Giavazzi, Charlotte Jacquemot, Sid Kouider

**Toddlers:** Katrin Skoruppa & Nivedita Mani

Isabelle Brunet, Emmanuel Dupoux, Anne-Caroline Fiévet, Kim Plunkett