PAROLE ET LANGAGE
UMR 6057 CNRS

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

Imitation in speech: Relationships between production and perception

Noël Nguyen

Laboratoire Parole et Langage

CNRS & Université d'Aix-Marseille

ILIKS Workshop

Aix-en-Provence, 25 mai 2011

General background

Current research on spoken communication seeks to better understand how spoken language is produced and perceived by humans in the context of its primary site of occurrence, i.e. social interaction.

Whereas earlier research has traditionally focused on laboratory speech, produced by single individuals, more recent work has provided evidence suggesting that the way in which language is used in everyday conversational exchanges, has an impact on how it is cognitively represented.

Usage-based models of language (Barlow & Kremmer, 2000; Bybee 2006; Couper-Kuhlen & Ford, 2004) very much emphasize the social dimension of language.

A main assumption is that the cognitive representations that are brought into play in the production and processing of spoken language do not entirely preexist to the interactions that may take place between talkers, but are rather subject to a co-construction process in which both interactants are engaged.

Convergence effects

In the course of a conversational interaction, the behavior of each talker can evolve with respect to that of the other talker in two opposite directions:

divergence $\longleftrightarrow$ convergence

Convergence effects are systematic and recurrent. They manifest themselves under many different forms: posture (ex. Shockley et al., 2003), head movements and facial expressions (ex. Estow et al., 2007; Sato & Yoshikawa, 2007) and, as regards speech, vocal intensity (Natale, 1975), pitch curve (Bosshardt et al., 1997; Gregory et al., 1993) and rate of speech (Giles et al., 1991).

These phenomena may facilitate conversational exchange by setting a common ground between speakers.

They may have the same effect as so-called alignment mechanisms, assumed to apply to linguistic representations at different levels between partners, in order for these partners to have a better joint understanding of what they are talking about (Garrod & Pickering, 2004).

PAROLE ET LANGAGE

(Very) long-term impact of imitation

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

While imitation occurs, by definition, within a social interaction, it has consequences for language that extend much beyond the temporal limits of that interaction.

Imitation is said to play a central role:

▶ In the acquisition of phonology, among the many aspects of language development (Studdert-Kennedy, 2002; Goldstein, 2003);

▶ In the emergence and evolution of human languages (de Boer, 2000).

In addition, imitation in humans embraces a domain that is much wider than that of language itself. It is central to the development of self in relationship with others (Piaget, Meltzoff).

Cerebral underpinnings of imitation

PAROLE ET
LANGAGE

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

The behavioral tendency shown by humans to imitate others may be connected at the brain level with the presence of mirror neurons (Rizzolatti et al., 2007).

These neurons may play a role in many different domains, from sensorimotor integration to the understanding of others' behavior (Arbib et al., 2006; Billard & Schaal, 2004, 2006; Fadiga et al., 2002; Gentilucci & Bernardis, 2007; Heiser et al., 2003 ; Hurley & Chater, 2005; Iacoboni et al., 1999; Jackson et al., 2006; Oztop et al., 2006; Rizzolatti & Craighero, 2007; Skipper et al., 2005; Watkins et al., 2003; Wilson et al., 2004. . .).

Consequences of imitation for how words are represented in long-term memory

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

Imitation in conversational interaction seems to affect the phonetic characteristics of the utterances produced by both speakers.

Previous studies (e.g. Pardo, 2006) have indeed shown that perceived similarity in pronunciation between talkers increases over the course of the interaction and persists beyond its conclusion.

This suggests that the representations associated with words in the mental lexicon for each talker, may dynamically evolve during conversation under the in?uence of the other talker's speech patterns, and retain the traces of that influence once the conversation has ended.

Thus, phonetic convergence has important implications for the current debate on how words are represented in the mental lexicon, and may be used as a probe for exploring the nature of these lexical representations.

PAROLE ET LANGAGE

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

Convergence effects appear consistent with the view that every word token leaves a trace in memory that contains fine-grained information about the particular circumstances in which this word token was produced or encountered.

This may include sensorimotor, semantic and pragmatic characteristics, but also indexical information about the talker's individual/social identity and the situation of occurrence.

Convergence effects seem difficult to reconcile with standard abstractionist models of speech production and speech processing (e.g. Stevens, 2002).

They provide support for the existence of an episodic-memory component (Goldinger, 1998), which may, however, be combined with abstract phonological representations, as in the hybrid approach, halfway between the abstractionist and exemplar theoretical viewpoints (Pierrehumbert, 2006).

Very little work has been carried out on potential convergence effects at the phonological level.

Everyday observations, however, suggest that these effects do take place, e.g. between talkers that speak different regional varieties of the same language. In such situations, each talker may tend to imitate part of the phonological patterns associated with the interlocutor's variety.

Phonological convergence towards the non-native variety may occur owing to the activation of the passive phonological knowledge that the speaker has acquired about that variety.

Whereas phonetic convergence may be more specifically tied to the particular word forms that are produced in the course of a conversational exchange, phonological convergence should apply across the board to all of the words sharing the same structural description (see Pierrehumbert's (2006) "phonological principle".

The SPIM Project

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

SPIM: Imitation in speech, from sensori-motor integration to the dynamics of conversational interaction

3-year project funded by the Agence nationale de la recherche

Three main partners:

- ▶ Laboratoire Parole et Langage, Aix-en-Provence
- ▶ Département Parole et Cognition, GIPSA-lab, Grenoble
- ▶ Laboratoire ICAR, Lyon

The main goal of this project is to gain a better understanding of the mechanisms that underlie imitation in speech. Our investigations combine data collected both in spontaneous conversational interactions and in a variety of experimental settings.

URL: spim.risc.cnrs.fr

One focus of interest in this project lies at the interface between phonetic and phonological convergence.

Specifically, we seek to determine to what extent phonetic convergence may set the ground for the formation of new abstract and generic phonological patterns.

Can imitative speech behavior in a given talker, tightly coordinated as it is with the other talker's own productions, have a more long-lasting and general impact on the talker's phonological knowledge?

We use imitation in speech as a means for examining the plasticity of phonological knowledge and its sensitivity to conversational interactions.

PAROLE ET
LANGAGE

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

Speech material

We seek to characterize convergence effects between talkers both within the same regional variety of French, and across regional varieties.

In the latter case:

▶ Talker 1: Southern French

▶ Talker 2: Northern French

Phonological / phonetic differences between Southern French and Northern French include:

▶ The presence/absence of a word-final open- vs close-mid vowel contrastive distinction

▶ The phonetic realization of back mid vowels.

PAROLE ET LANGAGE
UMR 6057 CNRS

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

Phonological / phonetic differences between Southern and Northern French

Northern French: contrastive distinctions between /e/-/ɛ/ and between /o/-/ɔ/

| été | [ete] | saute | [sot] |
| étais | [etɛ] | sotte | [sɔt] |

Southern French:

▶ only has close-mid /e/ and /o/

▶ [ɛ] and [ɔ] appear at the phonetic level but they are in complementary distribution with respect to the corresponding mid-high variants, according to a variant of the so-called loi de position (a mid-vowel phoneme is realized as close-mid in an open syllable and as open-mid in closed syllables and whenever the next syllable contains schwa, Durand, 1990).

| été | [ete] | saute | [sɔt(ə)] |
| étais | [ete] | sotte | [sɔt(ə)] |

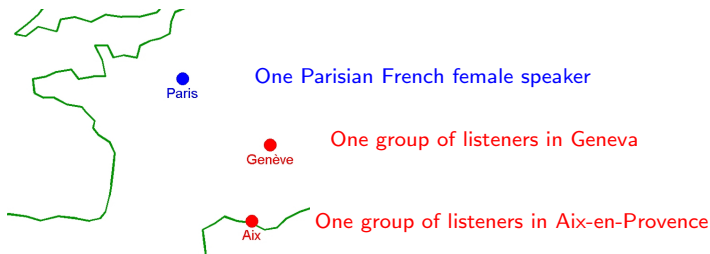# Processing a non-native regional variety: perceptual data

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

We examine how phonemic contrasts specific to one regional variety are perceived by listeners of another regional variety within the same language in a word recognition task.

More specifically, we explore how Southern French and Northern French listeners perceive /e/-/ɛ/ and /o/-/ɔ/ minimal word pairs.

One Parisian French female speaker

One group of listeners in Geneva

One group of listeners in Aix-en-Provence

The repetition priming paradigm

▶ Auditory lexical decision task

▶ Stimuli include words and non-words

▶ Test stimuli are presented twice

▶ Priming effect: Listeners are expected to respond faster to words
that have already been presented before

▶ Response patterns provide information about processing
mechanisms in real-time word recognition

PAROLE ET
LANGAGE

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

piquer	$\rightarrow$	RT1		feuquer	$\rightarrow$	RT1
gazon				gazon		
mulot				mulot		
crassue				crassue		
outil				outil		
robou				robou		
patin				patin		
piquer	$\rightarrow$	RT2		feuquer	$\rightarrow$	RT2

$RT2 < RT1$ $\qquad\qquad\qquad RT2 \approx RT1$

Same pairs and minimal pairs

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

same pair

piquer [pike]
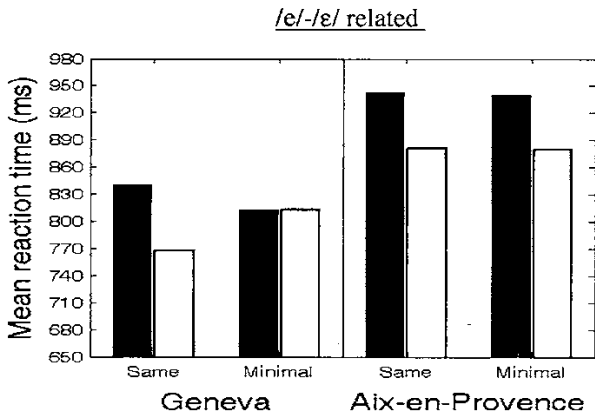gazon
mulot
crassue
outil
robou
patin
piquer [pike]

min. pair

piquer [pike]
gazon
mulot
crassue
outil
robou
patin
piquet [pikɛ]

PAROLE ET
LANGAGE
UMR 6057 CNRS

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

Observed response patterns

Aix listeners treated the 2nd member of /e/-/ɛ/ minimal pairs as a repetition of the 1st member. One potential interpretation is that these listeners perceptually assimilated both [e] and [ɛ] to a single phonemic category.

(Dufour, Nguyen & Frauenfelder, J. Acoust. Soc. Am., 2007)

One of the goals of the SPIM project relates to the potential impact of imitation on word recognition.

Does phonetic/phonological convergence towards the other speaker allow us to better understand that speaker?

It may be assumed that convergence contributes to making each speaker more attuned to the phonetic characteristics of words produced by the other speaker, via a perceptuo-motor loop.

We focused on the potential convergence effects that native speakers of Northern French may exhibit towards a native speaker of Southern French.

Material

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

The material was made up of a list of words that included 20 CVC test words, e.g. *dose*, with a back mid vowel expected to be produced

▶ as close-mid in Northern French, [doz]

▶ as open-mid in Southern French, [dɔz(ə)]

The entire list of words was produced by a native speaker of Southern French.

PAROLE ET
LANGAGE
UMR 6057 CNRS

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

Experimental design

18 speakers of Northern French were divided into two groups, the *test* group and the *control* group.

In the first stage of the experiment,

▶ The test group performed an *immediate shadowing task*: each participant had to repeat a list of words previously recorded by the Southern-French speaker

▶ the control group performed a *semantic categorization task*, which entailed no overt speech production, on the same list of words

In the second stage, all participants perform a word recognition task, on both the words they have already heard, and a list of nonwords.

Assuming that imitation facilitates later word recognition, we expect the test group's performance to be better than that of the control group.

Main findings

Acoustic analyses were carried out on the productions of the test group in the shadowing task.

These analyses revealed that most speakers in that group displayed a tendency to imitate the way in which words were produced by the Southern-French speaker.

Imitation took place in a systematic manner from the outset of the shadowing task for most speakers.

However, the lexical-decision experiment showed no facilitatory effect of imitation on word recognition. Reaction times to real words did not significantly vary depending on the subject group.

Thus, imitation of the non-native accent was found to be both rapid and systematic in shadowing but did not appear to facilitate later recognition of the critical words.
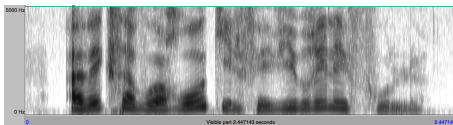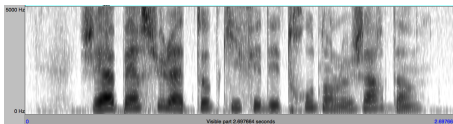
Northern participant

Pretest: Reading aloud

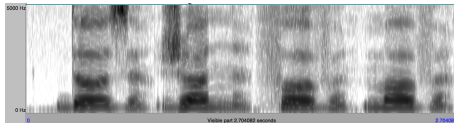*Avec sa voix* **rauque** *il fait vibrer les foules*

J'ai aperçu la **taupe** *qui fait des trous dans le jardin*

Northern participant

Test: Shadowing, target words produced by speaker of Southern French

dose . . . jaune . . . fauve . . . mauve

Conclusions

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

Our results reveal a systematic tendency for convergence towards the speaker of a non-native variety in speech production, in both a conversational and a laboratory setting.

To what extent convergence effects occur within the native variety is yet to be established. In a work in progress, Sato and colleagues find convergence effects for F_0 but not for vowel height (F_1 frequency) in a vowel repetition task in the native variety.

In word recognition, however, our data show that non-native contrastive distinctions may, in part, be abstracted away during lexical access.

In the recognition of spoken words, therefore, listeners exhibit a bias towards the phonemic repertoire of their native variety .

Conclusions (cont.)

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

In short, our data point to the representations associated with words in memory being both plastic (convergence across varieties in production) and stable (bias towards native variety in comprehension).

Our data are consistent with recent models of speech production and understanding that aim to bridge the gap between the exemplar-based and abstractionist approaches.

For example, Tuller and her coworkers (Tuller et al., 94; Tuller, 2004; see also Nguyen, Wauquier & Tuller, 2009) have proposed a model that uses concepts from the theory of nonlinear dynamical systems to account for the mechanisms involved in the categorization of speech sounds, and according to which there are two complementary aspects to speech perception.

Conclusions (cont.)

ILIKS Workshop
Aix, 25/05/11

State of the art

The SPIM Project

Speech material

Perception data

Production-
perception
links

On the one hand, speech perception is assumed to be a highly
context-dependent process sensitive to the detailed acoustic structure of the
speech input.

On the other hand, it is viewed as a non-linear dynamical system characterized
by a limited number of stable states, or attractors, which allow the system to
perform a discretization of perceptual space and which are associated with
abstract perceptual categories.

The recent development of so-called hybrid models (e.g., Hawkins, 2003,
2009; McLennan & Luce, 2005; Pierrehumbert, 2006) is also governed by the
assumption that detailed phonetic properties and abstract phonological
categories combine with each other in the representations associated with
words in memory.

Thank you

PAROLE ET LANGAGE
UMR 6057 CNRS

Nguyen, N. (2011). Representations of speech sound patterns in the speaker's brain: Insights from perception studies, in Cohn, A., Fougeron, C., & Huffman, M., eds, **Handbook of Laboratory Phonology** (Oxford Univ. Press), in press.

Brunellière, A., Dufour, S., & Nguyen, N. (2011). Regional differences in the listener's phonemic inventory affect semantic processing: A mismatch negativity (MMN) study, **Brain and Language** 117, 45–51.

Aubanel, V., & Nguyen, N. (2010). Automatic recognition of regional phonological variation in conversational interaction, **Speech Communication** 52, 577-586.

Dufour, S., Nguyen, N., & Frauenfelder, U. (2010). Does training on a phonemic contrast absent in the listener's dialect influence word recognition?, **Journal of the Acoustical Society of America Express Letters** 128 EL43-EL48.

Brunellière, A., Dufour, S., Nguyen, N., & Frauenfelder, U.H. (2009). Behavioral and electrophysiological evidence for the impact of regional variation on phoneme perception, **Cognition** 111, 390–396.

Dufour, S., Nguyen, N., & Frauenfelder, U. (2007). The perception of phonemic contrasts in a non-native dialect, **Journal of the Acoustical Society of America Express Letters** 121, EL131-EL136.