The International Journal of Speech, Language and the Law

IJSLL (PRINT) ISSN 1748-8885
IJSLL (ONLINE) ISSN 1748-8893

PhD Abstract

The influence of NATURE and NURTURE on speaker-specific parameters in twins' speech: acoustics, articulation and perception

Melanie Weirich

Research Assistant
Institut für Germanistische Sprachwissenschaft
Friedrich-Schiller-Universität Jena
Fürstengraben 30
07743 Jena
Germany

Awarding Institution: Humboldt-Universität zu Berlin, Germany
Date of Award: November 2011

KEYWORDS INTER-SPEAKER VARIABILITY, TWINS, ACOUSTICS, ARTICULATION, NATURE-NURTURE

The work presented in this dissertation set out to evaluate articulatory, acoustic and perceptual similarities and differences among monozygotic (MZ) and dizygotic (DZ) twins. In this respect, this dissertation contributes to the discussion of the impact of NATURE versus NURTURE on the production and perception of speech in general and on inter-speaker variability in particular. It is acknowledged that the sample is too small to offer a complete and generalisable account of the effect of nature/nurture on speech. Rather, it should be viewed as

Contact
email: melanie.weirich@uni-jena.de

IJSLL VOL 19.1 2012 119–122 doi : 10.1558/ijsll.v19i1.119

a case study, and the results could provide a stepping stone for future investigations. The aim of this study was to evaluate parameters that differ in the amount of inter-speaker variability between MZ twins (who have identical genes and similar physiology) and DZ twins (who share only about 50% of their genes and differ in their physiology). The assumption that MZ twin pairs reveal nearly identical physiological characteristics relevant in the speech production process (such as the size of the palate and the tongue or the palatal contour) while DZ twin pairs show significant differences regarding these parameters was checked and verified through several measurements and analyses. Furthermore, it was assumed that both groups of twins shared their social environment and that their speech acquisition process was oriented towards the same auditory goals. Thus, the aim of the study was to evaluate whether speaker-specific variability reflects *physiological differences* between speakers (NATURE) or is based on *learned variation due to social environmental influences* (NURTURE). In addition to complementing the general phonetic discussion of individuality in voice, the research also has theoretical and practical implications for the field of forensic phonetics.

Articulatory and acoustic data were recorded and analysed from four MZ twin pairs and three DZ twin pairs (native German speakers, same-sex, male and female, aged between 20 and 34 years). The speech material under investigation included vowels (/a, i, u/), sibilants (/s, ʃ/), and /aka/ sequences. The corner vowels were selected to get a representation of the whole vowel space. The sibilants and /aka/ sequences were deemed to be of special interest since a high degree of inter-speaker variability is discussed in the literature (for sibilants, see e.g. Dart 1998; Gordon, Barthmeier and Sands 2002; Toda 2006; for /aka/-gestures, see e.g. Mooshammer, Hoole and Kühnert 1995; Perrier, Payan, Zandipour and Perkell 2003). Additionally, a perception experiment was carried out to explore the perceived auditory similarity of the MZ and DZ twin pairs and sex-matched unrelated speakers. Acoustic analyses were performed to find acoustic correlates that explain the differences in perceived similarity.

The effect of NATURE should have a larger impact than the effect of NURTURE if a parameter differs more in DZ than in MZ twin pairs. Here, the influence of speaker- (or pair-) specific physiology on individual articulatory strategies and furthermore the role of somatosensory feedback were examined. If MZ and DZ twins showed the same amount of inter-speaker variability, then physiology (NATURE) played a secondary role to shared social environment, shared speech acquisition and auditory goals (NURTURE).

The results discussed in this dissertation point to the overall importance of NURTURE, shared social environment and the crucial role of auditory goals in speech. Nevertheless, NATURE revealed its impact in several aspects, too. In

particular, the potential influence of NATURE and/or NURTURE seemed not to be equally well reflected in all speech parameters. NURTURE is assumed to be intensified in parameters which correspond to certain linguistic units, e.g. to prominent syllables, whereas NATURE is more 'visible' in dynamic aspects of speech production, where a large degree of inter-speaker variability is found (cf. Kühnert and Nolan 1999; Rose 2002). In detail, three factors were identified that contribute to the relevance of NATURE versus NURTURE: 1) phoneme class (consonants vs vowels), 2) lexical stress (unstressed syllables vs stressed syllables) and 3) degree of coarticulation (i.e. the nature of the analysed item: transition vs target).

The first factor involves the *phoneme class* (1) and especially the differentiation between vowels and consonants. Since somatosensory feedback plays a larger role for consonants, which reveal a high amount of linguo-palatal contact (cf. Stone 1995), inter-individual differences in physiology were found to affect articulation more in sibilants and stops than in vowels. MZ twin pairs revealed greater similarity in the investigated parameters in stops and sibilants than DZ twin pairs, whereas for vowels no such differences could be found.

However, regarding vowels, a factor appeared that could intensify the impact of NATURE: *lexical stress* (2). It was hypothesized that syllables with lexical stress are more important regarding the communicative function of the speech signal than unstressed syllables. Syllables containing lexical stress are crucially dependent on learned auditory goals, whereas in unstressed syllables the speaker's individual physiology may be more relevant (cf. de Jong 1995). Indeed, a stronger impact of NATURE was found in parameters which are auditorily less salient: /i/ in an unstressed syllable (which is influenced by physiology) was more similar in MZ than in DZ twins, while for /i/ in a stressed syllable (which is strongly influenced by auditory goals) this was not the case.

The third significant factor deals with *coarticulation* (3). While it has been suggested that targets are learned entities that are influenced by and oriented towards shared auditory goals, the transitions between the targets are not controlled in the auditory domain. They are the consequence of the articulatory movements from one target to the next. The analyses conducted in this dissertation revealed that dynamic parameters of the speech signal – i.e. articulatory looping movements of the tongue back during /aka/ sequences, and F2 and F3 transitions in sibilant-schwa sequences – are more influenced by physiological constraints (NATURE) than static parameters, like articulatory target positions or acoustic characteristics in stable spectral regions. In this sense, transitions seem to mirror the individual properties of a speaker's vocal tract physiology and the biomechanical properties of the speaker's tongue muscles.

Results of the perception experiment confirmed that unrelated speakers are significantly easier to distinguish than twins, but zygosity has no effect on perceived similarity. Moreover, pair-specific auditory similarity appears in twins and unrelated speakers and can be explained by the acoustic parameters F_0, shimmer, jitter and HNR.

Thus, I concluded that idiosyncratic features are present even in the speech of twins and that both NATURE and NURTURE are crucial influencing factors in speaker-specific variability. However, the relative importance of the two factors is highly dependent on the specific characteristics of the investigated parameter.

References

Dart, S.N. (1998) Comparing French and English coronal consonant articulation. *Journal of Phonetics* 26: 71–94. http://dx.doi.org/10.1006/jpho.1997.0060

de Jong, K.J. (1995) The supraglottal articulation of prominence in English: linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America* 97: 491–504. http://dx.doi.org/10.1121/1.412275

Gordon, M., Barthmaier, P. and Sands, K. (2002) A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association* 32: 141–174. http://dx.doi.org/10.1017/S0025100302001020

Kühnert, B. and Nolan, F. (1999) The origin of coarticulation. In W.J. Hardcastle and N. Hewlett (eds.) *Coarticulation: Theory, Data and Techniques in Speech Production* 7–30. Cambridge: Cambridge University Press.

Mooshammer, C., Hoole, P. and Kühnert, B. (1995) On loops. *Journal of Phonetics* 23: 3–21. http://dx.doi.org/10.1016/S0095-4470(95)80029-8

Perrier, P., Payan, Y., Zandipour, M. and Perkell, J.S. (2003) Influences of tongue biomechanics on speech movements during the production of velar stop consonants: a modeling study. *Journal of the Acoustical Society of America* 114(3): 1582–1599.

Rose, P. (2002) *Forensic Speaker Identification*. London: Taylor & Francis.

Stone, M. (1995) How the tongue takes advantage of the palate during speech. In F. Bell-Berti and L. Raphael (eds.) *Producing Speech: Contemporary Issues* 143–153. New York: American Institute of Physics.

Toda, M. (2006) Deux stratégies articulatoires pour la réalisation du contraste acoustique des sibilantes /s/ et /ʃ/ en français. *Actes des XXVI ès Journées d'Étude de la Parole,* AFCP – IRISA – ISCA, Dinard, 65–68.