Proceedings of Meetings on Acoustics

Volume 6, 2009 http://asa.aip.org

157th Meeting
Acoustical Society of America
Portland, Oregon
18 - 22 May 2009
Session 4pSW: Speech Workshop

4pSW3. Voicing Profile of Polish and German Sonorants in Obstruent Clusters

Jagoda Sieczkowska* and Grzegorz Dogil

***Corresponding author's address: Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, Azenbergstrasse 12, Stuttgart, 70174, Baden-Württenberg, Germany, jagoda.sieczkowska@ims.uni-stuttgart.de**

Our study aims at defining and analyzing voicing profiles of Polish and German sonorants in particular at tendencies of Polish [r] and German [R] in obstruent clusters. Thanks to the use of digital phonetic tools (IMS German Festival synthesis system) which extracted voicing profiles as a frame-by-frame basis in continuous speech, we were able to study all phonotactically possible contexts of sonorants available in our speech corpuses. Our investigations conducted on Polish and German have shown that in both cases left-phoneme context is crucial in determining voicing probability of following sonorant.

Published by the Acoustical Society of America through the American Institute of Physics

INTRODUCTION

The former title of the presentation was *The acquisition of sonorant voicing: a case of L1 to L2 exemplars' transfer*. In the following article, instead of the acquisition process, we focus on the voicing profiles of Polish and German rhotics in obstruent clusters which is a part of our research project.

Languages employ feature [voice], which is part of the set of the universal distinctive features, when their consonants' types can be described in terms of presence or absence of voicing in the possible phonotactic contexts that are relevant for each language (Jessen, 1998). From the speech production perspective, voicing is defined as the presence of a periodic vocal fold vibration that produces a periodic excitation signal. Its acoustic consequences can be observed in the spectrum as the presence of low-frequency energy (Dogil, 2004; Lombardi, 1991; Möbius, 2004). Most studies in voicing focus on voice onset time (VOT), which is defined as the time interval between the stop release and the onset of voicing of the following vowel. Apart from the VOT analysis, a few other approaches to voicing analysis have been proposed: three category division into 'voiced', 'voiceless unaspirated' and 'voiceless aspirated' (Abercrombie, 1967) or binary divisions based on articulatory features (Chomsky & Halle, 1968). Most languages contrast in feature [voice] (Keating, 1979) which is usually examined in terms of obstruent voicing profile. In Polish, not only obstruents but also sonorants form oppositions in terms of voicing status: "Polish appears to display the common phenomenon of obstruent devoicing in word-final position, as well as voice assimilations in consonant clusters, both within words and at word junctures; additionally, sonorant consonants are devoiced in some positions (Gussmann, 1992)." In case of German, fonotactical rules say that sonorants cannot follow obstruents syllable-finally. Studies conducted by Möbius (2004) show sonorants in left-vocalic and voiceless obstruent context taking into account syllable boundaries and demonstrating their devoicing tendencies in the case of clusters voiceless obstruent plus sonorant. This study is focused on automatic analysis of temporal voicing values and their specification in context. Thus the description of Polish and German sonorants' voicing profiles is scalar rather than binary.

Production and cavity confirguration while producing sonorants includes spontaneous voicing which is why they are considered to be voiced (Chomsky & Halle, 1968; Gussmann, 2007). Spontaneous voicing is meant by 'narrowing the air passage to a point where the rate of flow is reduced below the critical value needed for the Bernoulli effect to take place' following Chomsky and Halle's description (Chomsky & Halle, 1968). "Sonorants are therefore articulated as voiced rather than voiceless unless some contrary gesture intervenes, such as spreading apart or tightly constricting the vocal folds so as to inhibit their vibration. Formally, spontaneous voicing can be expressed as the introduction into laryngeally unmarked sonorants of the glottal tension dimension (which by default implicates the gesture [slack]) (Iverson & Sang-Cheol, 2007)". Gussmann (1992) notes that "Polish sonorants are voiced except for a position between voiceless consonants or after a voiceless obstruent before pause", like in krtań [kr̥taŋ] *larynx*, wiatr [vjatr̥] *wind*, kadr [katr̥] *frame* or rytm [rɨtm̥] *rhythm*. Polish sonorants' devoicing depends also on the dialect and other socio-phonetic features. For example if the sonorant's preceding obstruent preserves its voicing the word final sonorant is voiced as well. In some cases a devoiced cluster might sound unnatural (like in wydm [wɨdm̥] *dune, gen. pl*), however, despite that, there are variants of sonorants that can only be devoiced, e.g. the ones in derivatives *–izm, -yzm* (marks-izm [markɕism̥] *Marxism*) (Gussmann, 2007). German sonorants can occur in obstruent and stop clusters in syllable-onset position, in right-fricative context as well as sonorant-sonorant clusters, which however never occur in onsets.

German syllable structure is based on the sonority hierarchy (with obstruents at the lowest sonority-scale place and vowels at the highest), according to which it can have either increasing or decreasing sonority from the peak to the end of the syllable and it allows maximally two consonants each in the onset and in the nucleus or coda position (Wiese, 1996). The regularities say that there can appear a certain number of consonants after a long vowel with only one consonant possible after a short vowel in the same monosyllabic word. In other words, Wiese (1996) claims that a long vowel can be treated equivalently to a diphthong, while short a vowel behaves in a way that it leaves place for an additional consonant to follow in the same syllable. It is also noted that according to German phonotactics "a non-obstruent is always placed inside an obstruent in the syllable, and while all non-obstruents can form a syllabic nucleus, an obstruent cannot (Wiese, 1996)." German r-sound has two main variants: [R] which is a uvular sound and [ɐ] which represents vocalized 'r' (Wiese, 1996). Different occurrences of [ɐ] and [R] in derivations of the same words are either syllabic or non syllabic like in grösser [gRø:.sɐ] *bigger, atr.function* versus grössere [gRø:.sə.Rə] *bigger, pl., pred.function* (Wiese, 1996). In the same studies Wiese has also demonstrated that [R] occurs mostly in onsets and sometimes after short vowels like in word Herr [hɛR] *man*. Since Final Devoicing in German is a phenomenon that applies only to obstruents in syllable-final position and German phonotactics does not allow any word-final sonorants with left-voiceless-obstruent context, in our study we examine sonorants in word-initial and mid-word position.

Turning back to the syllable, we now describe shortly the pattern for Polish syllable structure in the relation to sonorants' occurrence. The number of Polish syllables in the word is always the same as the number of the vowels in the word. Thus we can say that only the vowel can form a syllable nucleus. "The only other element that can appear in the rhyme is a sonorant consonant (Gussmann, 1992)." It has been also observed by Gussmann (1992), that a single sonorant which separates two vowels will always be assigned to the following syllable, like in domy [do.my] *house, pl.* In modern Polish sonorants are admitted in the rhyme and in the word-final clusters of consonants like in łga[rstw] *lie, gen.pl*. (Gussmann, 1992). Consonanants occurring in Polish word and syllable-initial positions are said to have 'peculiar clustering possibilities' (Gussmann, 1992), where there are cases in which a sonorant can occur before an obstruent like in [mkn]ąć *to speed* or in four-consonantal sequences in monosyllabic words like in [drgn]ąć *to shudder* (Gussmann, 1992). As devoicing concerns sonorants in word-final position, Gussmann (1992) argues that it is dependent on syllable structure and the disyllabification of the sonorant, because sonorants cannot be incorporated into a syllable structure that violates the Sonority Sequencing Principle. On the other hand, other studies, e.g. Rubach (1996, 2008) argued that it is not the syllable structure that licences devoicing but that the voice assimilation of the sonorants is governed by the linear adjacency of laryngeal nodes. It has been also pointed out in (Rubach, 1996) that rules of voicing in Polish interfere with acquisition of foreign languages by Poles. Common errors occur due to Final Devoicing (ex. [k] instead of [g] in *big*) and Spreading ([s] instead of [z] in *these people*) irrespective of their position in a prosodic structure (Rubach, 1996). Similarly Rubach (1996), claims that sonorants which violate sonority relations are transparent to Final Devoicing and to Regressive Assimilation (e.g. Jędrka [jent̬rka] *gen.sg*). Devoicing is also a matter of dialectal forms of Polish: Warsaw vs. Cracow form. Voice assimilation in obstruent clusters is the same in both dialects, voicing spreads in Cracow dialect, but not in Warsaw dialect [samɔχut rɔmana]. In our research we verify phonological hypotheses on sonorant voicing in Polish and German with the help of computational phonetic tools designed to operate on large phonetically labeled data.

METHOD

Voicing Profile

The method that we present differs from traditional acoustic measurements of voicing because the traditional approaches fail to provide sufficient information for sonorants occurring in different phonological environment. Thanks to the automatic tools which extract voicing profiles as a frame-by-frame basis in continuous speech, we can study all phonotactically possible contexts of sonorants. As shown in Figure 1, in the first step we used ESPS get_F0 tool to obtain frame by frame voicing probabilities at a sampling rate of 100 frames per second. After that get_F0 used an autocorrelation technique to calculate voicing probability. The value of those probabilities is always categorical (0 or 1).

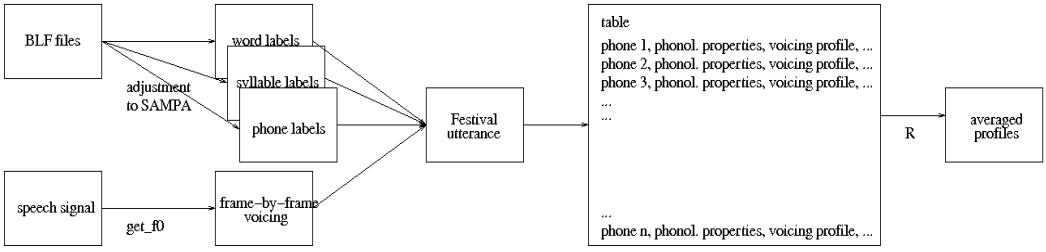

FIGURE 1: Illustration of methodological steps of the study.

Extraction of the features

We used the IMS German Festival synthesis system as it provides a wide range of information taken from the corpus about prosodic, syllable and segmental structure of the utterances (IMS German Festival Home Page, 2009). For each segment we extracted the following features: consonant type, place of articulation, voicing structure of the segment, its

preceding and following phone, word position, syllable structure, position in the syllable and voicing profile by extracting the voicing probability at 10%, 20%.... 90% of its duration. If the time of the phone of our interest happened to lie between two different probabilities' values, its value was obtained by interpolation. Using R (R project, 2001), the voicing profiles of phones in the selected contexts were averaged and plotted.

Databases

The analysis of sonorant consonants' profiles was conducted on two speech corporas. For Polish we used the BOSS corpora designed for the Polish version of the BOSS unit selection speech synthesis system (Demenko, 2008; Klabbers 2001). It comprises 115min (3249 utterances) of speech recorded by a professional speaker during several recording sessions supervised by an expert phonetician. The corpus consists of several databases: with most frequent consonant structures, diphones realized in 92 grammatically correct but semantically nonsense phrases, CVC triphones in non-sonorant voiced context and with various intonation patterns, CVC triphones in sonorant context and with various intonation patterns and utterances with 6000 most frequent Polish vocabulary items. 1109 sentences were recorded (Demenko et al., 2008). For German we used the MS corpus which was designed for the purpose of unit selection speech synthesis in the SmartKom project (SmartKom, 2003; Schweitzer et al., 2003). The core of the corpus comprises a full set of diphones and rich combinations of phones and contextual factors including segmental context, syllable structure and intonational factors. In the 160min of recorded speech there are 2601 sentences, which include 17489 words with 56434 consonants (Möbius, 2004).

RESULTS

In the study of Polish we decided to focus on the sonorant with most variable voicing – trilled [r], as our data show that other sonorants are categorically voiced in all contexts. After further specification of the phoneme context it has been observed that liquid [r] in word-final position preceded by voiceless stop like in word 'wiatr' [vjat̪r̥] *wind*, undergoes devoicing in almost all the exemplars in our corpus (Figure 2). It is also demonstrated that voicing profile of word-initial Polish sonorants in all-left-phoneme context show the same tendency– only liquid [r] undergoes devoicing in sizeable amount. Liquid [r] in word- initial position following voiceless stop, like in word 'rtęć' [rt͡ɛ̃ɕ] *mercury*, devoices [r] in more exemplars than in general all-phoneme-left context, similarly to word-final position, however here the amount of devoiced exemplars is higher and has a value of slightly more than 80%.

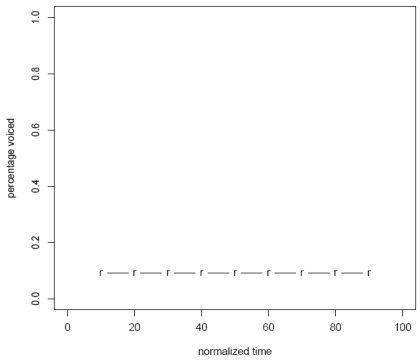

FIGURE 2*:* Voicing profile of Polish word-final [r] in voiceless stop left context.

Figure 3 (Möbius, 2004) shows voicing profile of German sonorants with left-voiceless obstruent context with and without syllable boundary. It has been observed that left context of the sonorant is the most important factor deciding about its voicing profile. All sonorants shown in the figure (including glide [j] which status is either a sonorant or a fricative or a non-syllabic vowel in German) undergo initial devoicing with the extent differing between sounds. In our

analysis we focus on the profile of [R] which tends to be voiceless through all its time duration when preceded by voiceless obstruent. Here again different realization possibilities of [R] have been observed (Möbius, 2004), which after observation of articulatory manners of the speaker, have been reported to be either velar or uvular voiced fricative, which in its devoiced variant is very close to [x]. It is noted however that devoicing effect in German may depend on presence vs. absence of syllable boundary like in word 'Stecknadel' [ʃtɛk.na:dəl] *pin* with syllable boundary and the word 'knapp'[knap] *tight* without syllable boundary. This pattern is observed for [j] and [R] in the biggest amount, contrary to the nasals.

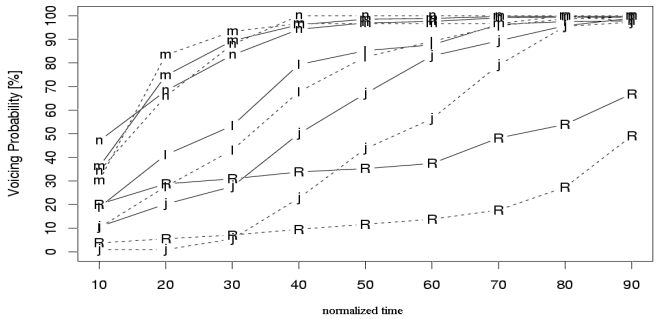

FIGURE 3 (Möbius, 2004)**:** Voicing profile of German sonorants with left-voiceless obstruent context with presence (solid lines) and absence (dashed lines) of the syllable boundary.

CONCLUSIONS

In this study we analyzed Polish word initial and final sonorants in general left context (all-phonemes) and following voiceless stops as well as German sonorants in initial and mid word position with the same left-voiceless-obstruent context. For Polish we observed that, almost exclusively, the sonorant [r] undergoes devoicing in those contexts out of the data we have analyzed. The most interesting is the case of [r] following voiceless stops, which is almost completely devoiced word and syllable-finally. Obstruent-liquid clusters are a peculiarity of the Polish language and have thus received many phonological descriptions (Cyran, 2005; Gussmann, 1992; Rubach 1996, 2008). The word-final liquid which follows an obstruent has to be accommodated into a phonological string by special abstract devices like extrametricality, segment-government or an empty nucleus. Such a special phonological status licenses phonetic behavior of the odd sonorant. Voicing profiles extraction based on our data-driven computational procedures demonstrate that both the phonological status and the prosodic position of the liquid influence its phonetic properties. However, final word position of the sonorant [r] governs its voicing profile more than its phonological environment - in the example of word initial and final [r] followed by voiceless stops we observed that the devoicing is much stronger than in the initial position, despite the same preceding phonological context. Studies conducted on German (Möbius, 2004) have shown that the probability of voicing is dependent on presence vs. absence of syllable boundary, which tends to weaken sonorants' devoicing preceded by voiceless obstruent. Thus both languages' sonorants voicing profile depends on syllable structure or positioning. In Polish it is more the word and syllable position that determines the amount of feature [voice] in [r], while in German it is the lack or the presence of the syllable boundary that occurs between voiceless obstruent and sonorant [R].

ACKNOWLEDGEMENTS

This study is part of the research project SFB 732 A2 funded by the German National Science Foundation (DFG).

REFERENCES

1. D. Abercrombie, 'Elements of General Phonetics'. Chicago Aldine Atherton, Inc. 1967. pp.56-84.
2. N. Chomsky and M. Halle, "The sound pattern of English'. Harper and Row. NY. 1968. pp. 298-302.
3. E. Cyran, "Sound pattern of Polish: phonotactic paradoxes at the right edge of words". Studies in Polish Linguistics. Vol 2, 2005. Pp. 61-89.

4. G. Demenko, J. Bachan, B. Möbius, K. Klessa, M. Szymański, S. Grocholewski, “Development and Evaluation of Polish Speech Corpus for Unit Selection Speech Synthesis Systems”. Interspeech Proceedings 2008.
5. G. Demenko, B. Möbius, K. Klessa, “The design of Polish Speech Corpus for Unit Selection Speech Synthesis.” In press (Language Technology vol.11). 2008.
6. G. Dogil et al. “Voicing”, Special Issue of Folia Linguistica XXXVIII/1-2. Mouton de Gruyter. Berlin. 2004.
7. E. Gussmann, “Resyllabification and Delinking: The Case of Polish Voicing”. Linguistic Inquiry 23, 1992. pp. 29-56.
8. E. Gussmann, “The Phonology of Polish”. Oxford University Press. 2007. pp.288-301.
9. IMS German Festival Home Page. http://www.cstr.ed.ac.uk/projects/festival/. 2009.
10. The International Phonetic Association (IPA) Homepage. http://www.arts.gla.ac.uk/ipa/ipa.html, accessed on 13 Apr 2008.
11. G. Iverson and A. Sang-Cheol, “English voicing in dimentional theory”. Language Sciences 29. 2007.
12. M. Jessen, “Phonetics and Phonology of Tense and Lax Obstruents in German”. John Benjamins Publishing Company. Amsterdam/Philadelphia. 1998. pp. 260-267.
13. P. Keating, ‘A phonetic study of a voicing contrast in Polish’. PhD Thesis. Brown University. 1979.
14. E. Klabbers and K. Stöber “Creation of Speech Corpora for the Multilingual Bonn Open Synthesis System.” ISCA ITRW. Scotland. 2001.
15. E. Klabbers, K. Stöber, R. Veldhuis, P. Wagner, S. Breuer, “Speech Synthesis Development Made Easy: The Bonn Open Synthesis System“. Eurospeech Conference. Scandinavia. 2001.
16. L. Lombardi, “Laryngeal Features and Laryngeal Neutralization, Ph.D. dissertation. University of Massachusetts, Amherst. 1991.
17. B. Möbius, “Corpus-based investigagtions on the phonetics of consonant voicing”. Folia Linguistica 38 (1-2), 2004.pp. 5-26.
18. R project. The R Project for Statistical Computing. (http://www.R-project.org/]. 2001.
19. J. Rubach, “Nonsyllabic Analysis of Voice Assimilation in Polish”. Linguistic Inquiry 27:69-100. 1996.
20. J. Rubach, “Prevocalic Fatihfulness”. Phonology, Vol.25 Number 3. Cambridge University Press. 2008.
21. A. Schweitzer, N. Braunchweiler, T. Klankert, B. Möbius and B. Säuberlich „Restricted Unlimited Domain Synthesis“. Proceedings of the European Conference on Speech Communication and Technology (Geneva).2003.pp.1321-1324.
22. Ch. Shih, B. Möbius, B. Narasimhan, „Contextual effects on consonant voicing profiles: a cross-linguistic study”. Proceedings of the 14th ICPhS Conference . San Francisco. 1999.
23. SmartKom, 2003. Das Leitprojekt SmartKom: Dialogische Mensch-Technik-Interaktion durch koordinierte Analyse und Generierung multiplier Modalitäten. (http://smartkom.dfki.de/start.html).
24. R. Wiese, „The phonology of German“. Calrendon Press. Oxford 1996. pp. 33-37, 252-265.