Predicting mutual intelligibility of Chinese dialects from objective linguistic distance measures

Tang, C.; Heuven, V.J. van

Predicting mutual intelligibility of Chinese dialects from objective linguistic distance measures

Tang Chaoju [1] & *Vincent J. van Heuven* [2]

[1]**Foreign Language School, Chongqing Jiaotong University**
Xuefu Road 66, Nan'an District, Nanping, 400074 Chongqing, P.R. China
chaoju.tang@gmail.com

[2] **Phonetics Laboratory, Leiden University Centre for Linguistics,**
PO Box 9515, 2300 RA Leiden, The Netherlands
v.j.j.p.van.heuven@hum.leidenuniv.nl

Abstract

This paper predicts the mutual intelligibility of 15 Chinese dialects from objective distance measures. Empirical mutual intelligibility measures were obtained from functional intelligibility tests at the sentence level from 15 listeners for each of 15 Chinese dialects. We computed various proximity measures on the basis of shared phonemes and tones in the sound inventories of the 15 dialects. Next, Levenshtein (string-edit) distance measures were computed on the 764 common syllabic units ('zi' in Pinyin, i.e., a meaningful character with a complete transcription of segments and tone) shared by the same 15 Chinese dialects in the Dialect Sound Database of Modern Chinese (compiled by the Chinese Academy of Social Sciences). Unweighed and perceptually weighed Levenshtein distance measures were computed. We also included objective similarity measures on the 15 dialects that have been published by Cheng (1997). The best single predictor of mutual intelligibility between a pair of dialects was the percentage of cognates shared between them (r^2 = .548). Including all predictors afforded a highly accurate prediction of mutual intelligibility (R^2 = .877). A very reasonable prediction is afforded if we just add the lexical frequency of finals (syllable rhymes) shared by a pair of dialects (R^2 = .612).

1. Introduction

Tang & Van Heuven (2007) collected from native listeners of 15 Chinese dialects judgments of linguistic similarity and intelligibility of these dialects. This enterprise yielded 225 combinations of speaker and listener dialects for which we reported scores for judged linguistic similarity and for judged intelligibility. We established that judged intelligibility can be predicted rather well from judged linguistic similarity (and vice versa) with r = 0.888.

Next, in Tang & Van Heuven (2008, 2009), we collected functional intelligibility scores for the same set of 225 combinations of speaker and listener dialects, using separate tests to target intelligibility at the isolated-word and at the sentence level. We then established, first of all, that these two functional intelligibility measures converged with r = 0.928; such convergence was expected since word intelligibility is a prerequisite to sentence intelligibility. Second, we wanted to know the extent to which functional intelligibility (the 'real thing') in the more recent papers could be predicted from the 'quick and dirty' judgment tests of our earlier work. If near-perfect prediction is possible, we will not have to apply cumbersome functional tests in the future, but may rely on the more convenient judgment tests. The results revealed that the correlation between the functional word and sentence intelligibility scores and the intelligibility judgment scores is good (r = 0.772 and 0.818, respectively) but not good enough to advocate the unqualified use of judgment testing as a more efficient substitute for functional testing.

1.1. Functional intelligibility at the sentence level

In the present paper we will concentrate on just one part of our data, viz. the functional intelligibility scores of Chinese dialects as established at the sentence level. The materials were adapted from the American Speech in Noise (SPIN) test developed to establish the extent of a patient's hearing loss (Kalikow et al. 1977). We only used the high-predictability sentences, in which the sentence-final target word is easier to understand as more of the preceding words are recognized, as in *She wore her broken arm in a sling* (target word underlined). A set of 60 SPIN sentences was translated into Standard Mandarin as well as in each of the following dialects: Beijing, Chengdu, Jinan, Xi'an, Taiyuan, Hankou (Mandarin dialects), Suzhou, Wenzhou (Wu dialects), Nanchang (Gan dialect), Meixian (Hakka dialect), Xiamen, Fuzhou, Chaozhou (Min dialects), Changsha (Xiang dialect), and Guangzhou (Yue dialect).

Groups of listeners (15 listeners for each of the 15 dialects) listened to (different) sentences in each of the 15 dialects and were instructed to write down the equivalent in their native dialect of the last word (two characters) in each sentence presented to them (for details see Tang 2009: chapter 4, Tang & Van Heuven 2009).

Thirty bi-dialectal consultants (a male-female couple for each dialect, all consultants were also fluent in Standard Mandarin) determined for each listener which target words were correctly translated into Mandarin (13,500 data points). Intelligibility scores were then computed for each combination of speaker and listener dialect, yielding a

15 × 15 = 225 cell matrix, as illustrated in Appendix 1. An agglomeration tree was generated from the intelligibility scores using average linking between groups (Appendix 2). The tree shows that the mutual intelligibility scores result in a plausible tree structure, such that the six Mandarin dialects and the nine Non-Mandarin (Southern) dialects end up in different main branches of the tree.

Mutual intelligibility was defined by Cheng (1997) as the mean of the intelligibility of speaker A for listener B and of speaker B for listener A. Obviously, if the intelligibility of A and B is not the same as that between B and A, averaging the AB and BA intelligibility scores eliminates the asymmetry. The averaging operation was performed on all pairs of contra-diagonal cells *i*, *j* and *j*, *i* in the 15 (speaker dialects) by 15 (listener dialects) = 225 cells in the score matrix we collected. We then deleted the redundant part of the matrices, keeping only the non-redundant lower triangle (without the main diagonal), and used the remaining 105 scores in the comparisons below.

1.2. Objective linguistic distance measures

Tang (2009: chapter 5) collected a large number of so-called objective measures, all of which contain some information on similarity between (pairs of) Chinese dialects. She computed structural similarity measures based on a simple comparison of the sound and tone inventories of the 15 dialects, with and without weighing the sound units for their lexical frequency. She also determined to what extent words in all pairs of dialects are pronounced the same, separately for segmental and tonal aspects. This work was based on lists of phonetic transcriptions of 764 words (basic morphemes) in each of the 15 dialects made available by the Chinese Academy of Social Sciences (CASS). She also copied from the literature published measures of structural similarity between all pairs of our 15 dialects (Cheng 1997), determined on a much larger list of 2,770 words (or rather concepts) occurring in the dialects. Among the various measures published by Cheng there is one that deserves special attention: this is the only measure we have for lexical similarity among the dialects (percent cognates shared); all other measures relate to differences in sound structure (vowels, consonants, tones). We would now like to know to what extent all these structural similarity measures impart the same information, and, even more importantly, if these allow us to predict the experimentally-based, functionally determined, mutual intelligibility scores between pairs of Chinese dialects. The present paper is an attempt to answer the various questions identified here.

2. The predictors

(i) Counts on sound inventory. The first group of predictors was based on simple counts on the phoneme inventories of the 15 dialects in our sample. The inventories of the 15 dialects were copied from the surveys provided by Yan (2006) and checked against the website maintained by Campbell (Campbell 2009, see http://www. glossika.com/ en/dict/faq.php#1) The lists of segmental sound symbols and tones are included in appendices 5.2-5.7 in Tang (2009). We then drew up lists containing all the different initials, nuclei, finals, codas, and tones across the set of 15 dialects. In each list we specified for each entry (in the rows) for each of the 15 dialects (in the columns) whether the particular sound or tone was or was not part of the inventory. When the sound was in the inventory, this was indicated by a '1', when it was absent from the inventory, a '0' was entered. On such data proximity matrices were generated. The proximity between two dialects can be used to predict the mutual intelligibility between them.

(ii) Lexical frequencies. A second, potentially more sophisticated, set of predictors was derived from the word lists contained in the dialect sound database of Modern Chinese compiled by the Institute of Linguistics of CASS (Chinese Academy of Social Sciences) (cf. Hou 1994, 2003). Henceforth, we will call this the CASS database. The list we used contains 764 morphemes in Modern Chinese. For each morpheme, the dialectal variant (or variants) in each of forty dialects is/are listed, including the 15 dialects of our sample. For each variant, a segmental and tonal transcription is digitally available. Segmental transcriptions are fairly narrow; tones are specified in terms of the 3-digit scheme proposed by Chao (1928). We split up the transcriptions into separate segmental and tonal representations, and made a further split in the segmental transcriptions in terms of onsets (initials), and finals (rhymes). The latter were further subdivided into vocalic nuclei (including glides) and codas. The frequencies of the various segmental parts and of the tones were then computed (between 0 and 764). The basic data look very much like the inventories examined in the preceding sections, with one important difference: whereas the inventories merely specify the presence ('1') or absence ('0') of an item in a dialect, the data now specify the frequency of an item in the list of 764 items. The frequency results were used to generate a proximity matrix for the 15 dialects.

(iii) Levenshtein distances. The Levenshtein distance (LD) is based on the smallest number of string operations (insertion, deletion, substitution) needed to convert the phonetic transcription of a word in language A to its counterpart in language B (or vice versa). LD has proven to be successful for measuring phonetic distances between Dutch dialects (Heeringa 2004), and successfully validated against perceived distances between pairs of Norwegian dialects (Gooskens & Heeringa 2004).

Again using the transcriptions in the CASS database of 764 common morphemes in each of our 15 dialects, we computed LD between all pairs of 15 dialects, once with and once without applying some perceptual weighing of sound differences.[1] In the unweighed LD, any difference between two sounds is considered of equal weight. When perceptual weighing was applied, we used the number of distinctive feature levels that differed between two sounds

[1] The LO4 software package can be downloaded from http://www.let.rug.nl/kleiweg/LO4/

as the weighing criterion. Here insertions and deletions were weighed at 50% of the maximum distance between either two consonants or between two vowels (for details of the weighing procedure see Appendices 5.15a and 5.15b in Tang 2009).

A problem in the case of Chinese dialects is that we have no way of knowing how tonal differences should be weighed against segmental differences. For this reason we decided to compute LD separately for the segmental and tonal properties of the morphemes. We will then later compare to what each of these domains contributes to intelligibility scores.

(iv) Published objective distance measures. Cheng (1997) measured the phonological affinity among the 15 dialects in our sample on the *Hanyu Fangyan Zihui [Word list of Chinese dialects]* (Beijing University 1989). The *Zihui* provides digital transcriptions of over 2,700 words across the dialects. Cheng's first measure is based on the correlation of the lexical frequencies of the initials only (470 different types). The second measure uses the lexical frequencies of the finals (rhyme portions of the syllables, 2770 different types). The third measure only considers the lexical frequencies of the tone transcriptions (133 different tone transcriptions). The fourth measure is based on the segmental transcription of the initials and finals combined (470 initials + 2770 finals = 3240 different transcriptions). The fifth and last measure is the combination of the previous one plus the 133 tone transcriptions (3373 different transcriptions).

Two more objective distance measures were copied from Cheng (1997). We call these the Phonological Correspondence Index (PCI) and the Lexical Similarity Index (LSI). The PCI is a measure that expresses the complexity of the rule system that is needed to convert phonemic transcriptions (including tones) in dialect A to their cognate form in language B. The more complex the rule system, the larger is the distance between dialects A and B. Note that this is the only measure in our study that is not symmetrical: the rule set that converts A to B may be more or less complex than the set that converts forms from B to A (for details, see Cheng 1997). We transformed the asymmetrical PCI distance matrix to a symmetrical version as explained above. The symmetrical distances were used to predict the function mutual intelligibility scores. LSI was conceptually defined by Cheng (1997) on the *Zihui* word list as the percentage of cognates shared by two dialects. This is a symmetrical measure. Obviously, the larger the percentage of shared cognates the easier it should be for a speaker of dialect A to be understood by a listener of dialect B (and vice versa). This then is our last single predictor of mutual intelligibility.

3. Correlation and regression analyses

3.1. Single predictors of mutual intelligibility

The raw correlation coefficients between each of 27 objective linguistic similarity measures and the functional sentence intelligibility scores are included in Appendix 2. We will now determine the best, and most promising, single linguistic distance measures as predictors of mutual intelligibility of our Chinese dialects in each of five types of data as explained above: (i) sound inventories, (ii) lexical frequencies of similar sound units derived from the CASS transcriptions, (iii) string distance measures (Levenshtein) determined on the same collection of transcriptions, (iv) lexical frequencies of phonological units published by Cheng (1997), and (v) overall measures of lexical and phonological similarity published by Cheng (1997).

Within the similarity measures based on the sound inventories, finals, and especially coda elements (rather than vocalic nuclei) shared between dialects provide the best predictors of functional intelligibility (r-values around .500). Tones shared in the inventories are intermediate (around .400), and least successful predictors are shared initials (onsets) with r-values on the order of r = .250 (marginally significant).

The distance measures we derived ourselves from our lists of sound inventories in the 15 dialects reflect the same tendencies that were apparent in Cheng (1997). Again, the best correlations are found for shared finals (codas rather than nuclei), whilst shared initials (onsets) and tones are poorer predictors.

Also, when we consider the distance measures computed on the lexical frequencies of the sound units in the CASS transcriptions of 764 basic morphemes, we find the best (but not good) correlation for shared finals (r-values around .425), slightly poorer correlations for onsets, nuclei and codas (r-values between .360 and .400) and the poorest correlation for tones (around r = .220). Distance measures based on string-edit procedures correlate least with functional intelligibility scores (insignificant or marginally significant r-values between .038 and .326).

We now come to the simpler types of measures published by Cheng (1997). Among this group of objective distance measures the shared finals stands out with r-values around r = .720. Correlation coefficients for other phonological units are poorer, and no correlation at all is obtained for shared tones.

Much better predictions are obtained from the more comprehensive measures in Cheng (1997). Both the lexical (LSI) and the phonological (PCI) affinity correlate with word and sentence intelligibility with r-values of .740 and .772, respectively. We also note that the intercorrelation between lexical (LSI) and phonological (PCI) similarity is still low enough (r = .761) to make multiple prediction a worthwhile undertaking (§ 3.2).

3.2. Predicting mutual intelligibility by multiple regression

We will now attempt multiple regression analyses for the functional intelligibility scores. Unfortunately, LSI data were not available for the Mandarin dialects Taiyuan and Hankou. Therefore all multiple regression analyses were done on a reduced number of dialect pairs, i.e. 78 (instead of 105). The results of these analyses are presented in

Table 1, separately for predictors that were entered simultaneously, and for stepwise solutions.

Table 1. Results of Multiple Regression Analyses, predicting functional sentence intelligibility scores from non-compound objective measures of linguistic distance. CC: Data from Cheng (1997), Inv: our own data on sound inventories of Chinese dialects, CA: lexical frequencies based on the Chinese Academy of Social Sciences database. In the stepwise analysis the predictors were entered in the order listed (based on the highest partial correlation with the criterion variable); the R^2 values are cumulative. The absolute values of the beta weights indicate the relative importance of a predictor.

Simultaneous entry				Stepwise entry		
Predictors	R^2	β		Predictors	R^2	β
CC_LSI		.571		CC_LSI	.548	.621
CC_Finals		.278		CC_Finals	.612	.405
Inv_Tones		–.612		CA_Onsets	.680	.663
Inv_Initials		–.410		Inv_Finals	.725	–.498
CA_Onsets		.696		Inv_Tones	.759	–.646
CA_Tones		.481		CA_Tones	.816	.475
				CA_Finals	.846	.621
All	.877			Leven_weight	.855	.101

With simultaneous entry of all predictors we obtain a high R^2 value of .877 for sentence intelligibility, at least when all (non-compound) predictors are included. However, only six objective distance measures make a significant contribution to the prediction of sentence intelligibility. Note that PCI, which was the single most successful predictor of mutual intelligibility, plays no role in the multiple prediction. Its raw correlation with the criterion within the reduced set of 78 dialect pairs is lower than that of LSI. Subsequent partial correlations are always better for other predictors than PCI.

When we attempt stepwise entry of predictors, an R^2 value of .855 is found for sentence intelligibility with eight predictors. The first two predictors (CC_LSI and CC_Finals) are the same as in the simultaneous-entry solution, with roughly the same beta weights but from the third predictor onwards the results diverge. By and large, these results indicate that a fairly good prediction of sentence intelligibility can be obtained (R^2 = .612) from just two predictors, one that covers lexical distance (percent cognates shared) and one that covers phonological distance, i.e. lexical frequency of finals (syllable rhymes) in Cheng's (1997) count based on the 2,270 item *Zihui* word list.

4. Conclusion and discussion

The best prediction of mutual intelligibility between two Chinese dialects (within our sample of 15) by a single objective measure of linguistic distance is afforded by Cheng's (1997) Lexical Similarity Index (LSI), which is basically the percentage of cognates shared between the two dialects. This measure by itself accounts for 58% of the variance in the mutual intelligibility scores found for the 105 combinations of pairs of dialects in our sample. Interestingly, when we try to improve the prediction of mutual intelligibility, the most useful contribution is made by another objective measure computed by Cheng (1997). The objective measures that we computed ourselves correlate more poorly with the criterion than Cheng's most successful measures. We assume that the superiority of Cheng's measures is caused by the fact that he computed them on a much larger word list that ours, so that Cheng's measures have better coverage of our stimulus sentences. Ideally we should compute objective linguistic distance measures on the specific lexical materials that we used in our 15 (dialects) × 60 stimulus sentences. Unfortunately such an enterprise was beyond the scope of the present paper.

5. References

[1] Cheng, Chin-Chuan, 1997. Measuring relationship among dialects: DOC and related resources. *Computational Linguistics & Chinese Language Processing* 2, 41-72.

[2] Beijing University 1989. *Hanyu fangyin zihui di'er ban [Chinese Dialect Character Pronunciation List, second edition]*. Beijing: Wenzi Gaige Chubanshe (Philology Reform Publish house).

[3] Beijing University 1995. *Hanyu fangyin cihui di'er ban [Chinese Dialect word List, second edition]*. Beijing: Wenzi Gaige Chubanshe (Philology Reform Publish house).

[4] Chao, Yuan-Ren 1928. *Studies in the Modern Wu Dialects*. Tsinghua College Research Institute Monograph 4, Beijing.

[5] Gooskens, C., W. Heeringa, 2004. Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. *Language Variation and Change* 16, 189-207.

[6] Heeringa, W., 2004. *Measuring dialect pronunciation differences using Levenshtein distance*. Doctoral dissertation, Groningen University.

[7] Hou, Jinyi 1994. *Xiandai hanyu fangyan yiku [The sound databank of Modern Chinese]*. Shanghai: Shanghai Education Press.

[8] Hou, Jinyi 2003. *Xiandai hanyu fangyan yiku [The sound databank of Modern Chinese]* (CD-ROM version, in Chinese). Shanghai: Shanghai Education Press.

[9] Kalikow, D.N., K.N. Stevens, L.L. Elliott 1977. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America* 61, 1337-1351.

[10] Tang, Chaoju 2009. *Mutual Intelligibility of Chinese dialects: An experimental approach*. LOT dissertation series nr. 228. Utrecht: LOT.

[11] Tang, Chaoju, V. J. van Heuven 2007. Mutual intelligibility and similarity of Chinese dialects, in B. Los, M. van Koppen (eds.) *Linguistics in the Netherlands 2007*. Amsterdam: John Benjamins, 223-234.

[12] Tang, Chaoju, V. J. van Heuven, 2008. Mutual intelligibility of Chinese dialects tested functionally, in M. van Koppen, B. Botma (eds.) *Linguistics in the Netherlands 2008*. Amsterdam: John Benjamins, 145-156.

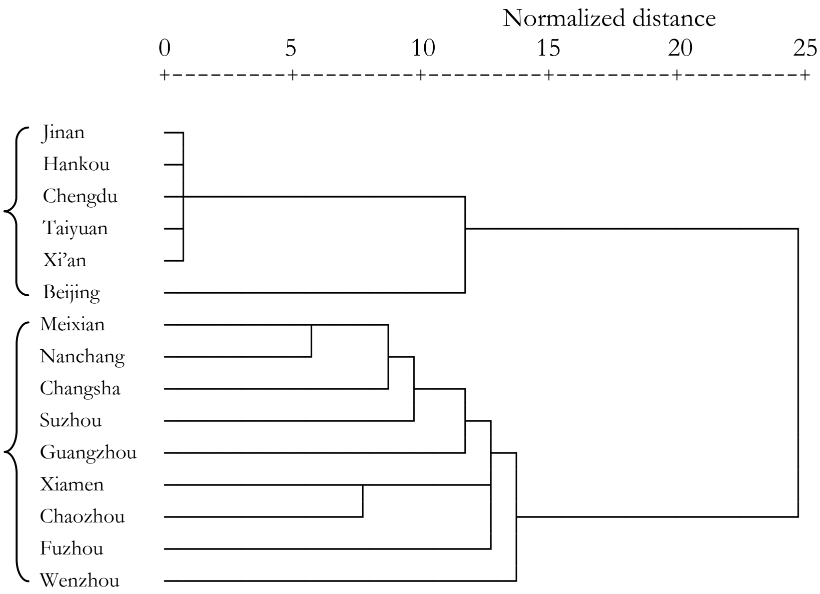
[13] Tang, Chaoju, V.J. van Heuven 2009. Mutual intelligibility of Chinese dialects experimentally tested. *Lingua* 119, 709-732.

[14] Yan, M. Mian 2006. *Introduction to Chinese Dialectology*. LINCOM Studies in Asian Linguistics. München: LINCOM.

Appendix 1. Percent correctly translated target words in sentences broken down by 15 speaker dialects and 15 listener dialects. Each mean is based on 60 responses (each of 60 sentence-final words is heard once, with 4 different words per dialect for each of 15 listeners). The total number of responses is 225 × 60 = 13,500. Double lines separate Mandarin from non-Mandarin dialects.

Speaker dialect (down)	Listener dialect(across)															
	Suzhou	Wenzhou	Guangzhou	Xiamen	Fuzhou	Chaozhou	Meixian	Nanchang	Changsha	Taiyuan	Beijing	Jinan	Hankou	Chengdu	Xi'an	Mean
Suzhou	**77**	7	5	18	13	5	7	13	13	20	5	18	15	15	7	16
Wenzhou	5	**93**	5	12	3	2	7	10	2	7	2	10	8	7	2	10
Guangzhou	5	7	**92**	10	20	25	55	22	13	7	3	22	8	17	7	21
Xiamen	13	5	8	**97**	23	28	13	18	13	3	5	15	7	17	8	18
Fuzhou	3	3	2	17	**92**	7	3	8	5	0	0	7	2	0	3	10
Chaozhou	7	0	3	52	13	**98**	3	12	3	7	2	13	10	3	5	15
Meixian	13	2	12	28	17	20	**70**	25	18	10	3	25	15	25	8	19
Nanchang	28	13	20	25	27	17	33	**50**	32	35	18	53	43	37	23	30
Changsha	12	3	8	23	17	3	17	25	**93**	13	13	38	53	28	2	23
Taiyuan	63	35	45	63	57	25	55	68	68	**73**	77	92	92	85	73	65
Beijing	87	62	90	90	93	60	80	78	92	90	**98**	98	97	98	93	87
Jinan	52	27	32	48	48	15	40	60	70	75	77	**97**	83	82	67	58
Hankou	48	32	32	52	53	27	45	53	62	58	67	95	**100**	73	65	57
Chengdu	47	22	40	48	72	27	48	58	62	65	62	98	95	**95**	68	60
Xi'an	53	33	50	58	57	30	57	58	63	68	58	82	78	70	**67**	59
Mean	34	22	30	43	40	26	36	37	41	35	33	51	47	43	33	

Appendix 2. Dendrogram (using average linking between groups) and Euclidean distance measures based on sentence-level intelligibility scores obtained for all 225 combinations of 15 speaker and 15 listener dialects. Note that the tree correctly reflects the primary split of the 15 dialects into a Mandarin and Non-Mandarin (Southern) group – as indicated by the braces.



Appendix 3 Correlation matrix between subjective and objective measures. CC: data from Cheng (1997), Inv: sound inventories of Chinese dialects, CA: lexical frequencies based on the CASS database. LD: Levenshtein distance. Bolded coefficients are significant ($p < .01$).

Variables	CC_init	CC_final	CC_init_itfin	CC_in_fin_tone	CC_tone	CC_LAI	CC_PCI	Inv_init	Inv_nuc	Inv_coda	Inv_tone	Inv_final	Inv_init_coda	Inv_init_final	Inv_in_fin_tone	CA_init	CA_final	CA_coda	CA_tone	CA_nuc	CA_init_final	CA_in_fin_tone	Lev_unw	Lev_weight	Lev_tone	Lev_tone_change	Tone_weighed
CC_final	**0.52**																										
CC_init_final	**0.95**	**0.76**																									
CC_in_fin_tone	0.19	0.23	0.26																								
CC_tone	0.03	0.10	0.09	0.99																							
CC_LAI	**0.45**	**0.60**	0.53	0.26	0.14																						
CC_PCI	**0.58**	**0.86**	0.74	0.29	0.17	**0.76**																					
Inv_init	**-0.50**	**-0.34**	-0.52	-0.37	-0.29	-0.27	**-0.44**																				
Inv_nuc	-0.05	**-0.38**	-0.16	0.03	0.06	-0.30	**-0.36**	0.23																			
Inv_coda	**-0.34**	**-0.43**	-0.40	-0.22	-0.16	**-0.64**	**-0.56**	**0.29**	**0.36**																		
Inv_tone	-0.20	**-0.39**	-0.29	-0.22	-0.17	-0.29	**-0.33**	0.13	0.22	0.22																	
Inv_final	-0.22	**-0.51**	-0.33	0.05	0.11	**-0.50**	**-0.53**	0.04	**0.75**	**0.59**	**0.29**																
Inv_init_coda	**-0.55**	**-0.48**	-0.60	-0.39	-0.30	**-0.51**	**-0.62**	**0.90**	**0.36**	**0.67**	0.21	**0.32**															
Inv_init_final	**-0.29**	**-0.56**	-0.40	-0.01	0.06	**-0.54**	**-0.59**	0.18	**0.76**	**0.62**	**0.31**	**0.99**	**0.44**														
Inv_in_fin_tone	**-0.30**	**-0.58**	-0.42	-0.03	0.04	**-0.55**	**-0.60**	0.18	**0.76**	**0.62**	**0.37**	**0.99**	**0.44**	**1.00**													
CA_init	**-0.76**	**-0.59**	-0.79	-0.22	-0.09	**-0.59**	**-0.68**	**0.69**	**0.26**	**0.40**	0.22	**0.33**	**0.73**	**0.42**	**0.43**												
CA_final	**-0.36**	**-0.66**	-0.51	-0.32	-0.24	**-0.45**	**-0.60**	**0.34**	**0.49**	**0.50**	**0.40**	**0.62**	**0.51**	**0.66**	**0.67**	**0.44**											
CA_coda	**-0.29**	**-0.40**	-0.35	-0.14	-0.08	**-0.41**	**-0.48**	0.25	**0.33**	**0.57**	**0.28**	**0.49**	**0.45**	**0.51**	**0.52**	**0.27**	**0.69**										
CA_tone	0.05	0.08	0.05	-0.18	-0.19	0.15	0.15	0.12	-0.02	-0.10	**0.69**	-0.15	0.04	-0.14	-0.08	-0.02	0.02	0.04									
CA_nuc	**-0.32**	**-0.63**	-0.47	-0.38	-0.31	**-0.31**	**-0.53**	**0.37**	**0.51**	**0.38**	**0.42**	**0.52**	**0.48**	**0.57**	**0.58**	**0.43**	**0.88**	**0.41**	0.04								
CA_init_fin	**-0.62**	**-0.75**	-0.74	-0.33	-0.21	**-0.62**	**-0.76**	**0.57**	**0.47**	**0.54**	**0.38**	**0.58**	**0.71**	**0.65**	**0.67**	**0.80**	**0.89**	**0.60**	0.01	**0.81**							
CA_ons_fin_tone	**-0.45**	**-0.53**	-0.54	-0.36	-0.28	**-0.37**	**-0.49**	**0.52**	**0.35**	**0.36**	**0.72**	**0.35**	**0.58**	**0.41**	**0.46**	**0.61**	**0.70**	**0.49**	**0.63**	**0.65**	**0.78**						
LD_unweighed	-0.15	**-0.34**	-0.23	-0.16	-0.13	**-0.32**	**-0.33**	0.14	0.23	0.18	**0.28**	**0.27**	0.21	**0.29**	**0.30**	0.21	**0.38**	**0.27**	0.00	**0.31**	**0.36**	**0.27**					
LD_weighed	-0.09	**-0.27**	-0.16	-0.13	-0.11	-0.27	**-0.25**	0.11	0.20	0.15	0.24	0.22	0.17	0.24	0.25	0.16	**0.34**	**0.26**	0.00	**0.26**	**0.31**	0.23	**0.98**				
LD_tone	0.18	0.06	0.15	0.06	0.04	0.21	0.08	-0.07	0.04	-0.10	0.13	0.00	-0.11	-0.01	0.00	-0.14	-0.08	-0.02	**0.31**	-0.04	-0.14	0.09	-0.22	-0.24			
LD_tone_change	0.01	0.02	0.01	0.12	0.12	0.17	0.02	0.03	0.02	-0.12	**-0.41**	-0.06	-0.03	-0.05	-0.08	0.05	-0.09	-0.16	**-0.37**	-0.04	-0.03	**-0.26**	-0.10	-0.10	0.06		
Tone_weighed	-0.03	0.04	-0.03	-0.10	-0.09	0.15	-0.05	0.18	0.06	0.06	0.04	0.07	0.14	0.08	0.08	0.21	0.20	0.15	0.13	0.20	0.23	**0.26**	0.00	0.03	**0.35**	**0.31**	
F_sentsence	**0.44**	**0.71**	0.56	0.06	-0.04	**0.74**	**0.77**	**-0.26**	**-0.33**	**-0.45**	**-0.44**	**-0.48**	**-0.43**	**-0.52**	**-0.54**	**-0.43**	**-0.42**	**-0.36**	0.15	**-0.39**	**-0.50**	**-0.27**	**-0.27**	-0.19	0.10	0.15	0.14