The relation between global pitch range and gestures in a story-telling task

M. Grazia Busà [1]*, Sara Brugnerotto* [1]

[1] DiSLL, University of Padova, Padova, Italy

Mariagrazia.Busa@unipd.it, Sara.Brugnerotto@studenti.unipd.it

Abstract

Anecdotal evidence suggests that both pitch range and gestures contribute to the perception of speakers' liveliness in speech. However, the relation between speakers' pitch range and gestures has received little attention. It is possible that variations in pitch range might be accompanied by variations in gestures, and vice versa. In second language speech, the relation between pitch range and gestures might also be affected by speakers' difficulty in speaking the L2. In this pilot study we compare global pitch range and gesture rate in the speech of 3 native Italian speakers, telling the same story once in Italian and twice in English as part of an in-class oral presentation task. The hypothesis tested is that contextual factors, such as speakers' nervousness with the task, cause speakers to use narrow pitch range and limited gestures; a greater ease with the task, due to its repetition, cause speakers to use a wider pitch range and more gestures. This experimental hypothesis is partially confirmed by the results of this study.

Index Terms: pitch range variation, gesture rate, story telling, English L2, Italian L1

1. Introduction

One of the goals of public speaking classes is to teach students to use a 'lively' voice when delivering a speech. This means that students should speak with a voice that varies in intonation, rhythm and volume. This is because by varying intonation, rhythm and volume speakers can emphasize important points of their discourse and deemphasize others, and thus help listeners follow the information flow. In other words, variation in speech helps listeners maintain their focus on the speaker's message and not wander away [1, 2].

In addition to voice, public speaking classes emphasize the importance of body language in discourse: students are told to maintain an open body position and to use gaze and gestures to highlight parts of speech. This contributes to maintaining the listeners' attention by providing them with a visual channel, in addition to the audio channel, that helps them follow the information flow.

For second-language learners, speaking in public involves planning thoughts, discourse structure and words, together with intonation and gestures, in a language that is not their own. This results in a very heavy cognitive load that may impair one or all levels of output: linguistic, prosodic, and gestural. As a result, second-language learners' delivery of speeches in public may appear incongruent or tedious, with an effect on the successful outcome of their presentations. However, in L2 as in L1, performance can be improved through preparation and rehearsal, which can contribute to reducing the contextual factors, such as nervousness, that affect speakers' congruence and delivery.

The worldwide success of public speaking classes shows that students can –in fact– learn to modify their voice and body language habits in discourse, and give oral presentations that are effective in holding the audience' attention.

However, though the dynamics of successful speaking attract the interest of many, there is a lack of scientific research focusing on the quantitative measurements of performance.

This paper reports on a preliminary study aimed at investigating how contextual effects, such as nervousness for a speech delivery, may affect speakers' use of pitch range and gestures. This is done by presenting an investigation of the global pitch range and gestural characteristics of 3 Italian speakers of English engaged in a story-telling task in Italian and English.

2. Pitch range, gestures and common ground

It is known that in most languages meaning and emphasis are created by means of variations of the fundamental frequency (or F_0) of the human voice. The range over which these variations may occur is called pitch (or F_0) range. Typically, a voice that is heavily inflected, that is, has a wide pitch range, will sound animated; a voice that has a narrow pitch range will sound monotone. Thus, pitch range has been used as a measure of speaker's perceived liveliness [1, 2, 3] –though the use and interpretation of pitch range may vary depending on language [3, 4, 5] and sociocultural/ sociophonetic factors [6].

It has been suggested that L2 speech may be characterised by limited pitch variation and a narrower pitch range than L1 speech [1, 2, 3, 4, 8, 9, 10, 11]. It is possible, in fact, that prosodic information is processed differently by native and non-native speakers because of their different levels of competence in the L1/L2. For example, as suggested by [7], non-native speakers may rely more on segmental, as opposed to prosodic, information to get their meanings across, given the fact that they lack the amount of extra-linguistic knowledge that native speakers can rely on when communicating. Differences in pitch range in L1 and L2 may also be more conspicuous in particular speaking styles, such as formal presentations [1, 2, 12], during which non-native speakers may be particularly focussed on getting their meanings across, at the expense of prosody.

A framework for measuring global pitch range cross-linguistically was first established by Ladd [13], then elaborated by Patterson [14], and finally by Mennen et al. [3; 4]. Within this framework, a number of measures are used to quantify differences in pitch level (i.e., the speaker's overall pitch height or register) and pitch span (i.e., the speaker's range of frequencies in a speech sample). These include F_0 max, min, mean and median, as well as linguistic measures, linked to specific linguistically-defined landmarks in the F0 contour.

A different measure of pitch range was used by Hincks [1, 2] to compare speakers' liveliness over long stretches of speech. Hincks looked at the normalized standard deviation of F_0, and found that a value of pitch variation, which she called pitch variation quotient (PVQ), strongly correlates with perceived speakers' liveliness, though only weakly with speakers' proficiency level. Pitch variation appeared to be a

stronger perceptual cue to liveliness in male speech than in female speech. She concluded that pitch variation may not be the *only* measure of speakers' liveliness (rhythm and intensity being also measures of liveliness), but it is certainly an important one.

Research has shown that speech and gestures are interconnected [e.g., 15, 16]. According to McNeil [17, 18], speech and gestures are synchronous at the semantic level, as they are co-expressive of the same underlying meaning; at the pragmatic level, as they co-occur to express the same pragmatic function; and at the phonological level, as gestures are temporally coordinated with the phonology of the utterances.

A number of studies have examined the relationship of prosody and gestures, focussing in particular on the investigation of the temporal alignment of gestures with prosodic prominence [e.g., 19, 20, 21, 22, 23]. Evidence has been found that gestures are coordinated with prosodic stress, but there is little consensus as to how exactly gestures are aligned with prominent parts of speech [e.g., 24, 25, 26, 27, 28, 29]. Beat gestures might have a stronger influence on speech production than representational gestures [30]. It is possible that some gestures have an effect on the perception of speech prominence. For example, the realization of a visual beat in association with a prosodically prominent word has an effect on the acoustic realization of the word, and causes that word to be perceived as more prominent than the neighboring words [30].

While research has focussed on the synchronization of gestures with prosodic prominence, the relationship between speakers' global pitch range and gestures has received little attention. Anecdotal evidence suggests that there might be a relation between the amount of pitch variation in speakers' speech and the extent to which speakers gesture when they speak. In fact, it is highly likely that speakers convey paralinguistic meanings through their voices as well as through their gestures.

Co-speech gestures seem to fulfill a number of functions, and may in fact be multifunctional [reviewed in 31, 32, 33]. Gestures have been shown to facilitate speakers' cognitive processes during speech production; for example, they seem to help speakers conceptualize, retrieve lexical items, manage cognitive loads, organize information into syntactic constituents. Gestures also seem to be planned and produced with the addressee's needs in mind, and so play a role in communication. For example, speakers produce more and larger gestures when they see their interlocutor(s), than when they do not (e.g., when they are talking over the phone) [34]. Speakers' gestures are also affected by common ground, that is the amount of knowledge that is shared between the participants in a spoken interaction. It has been shown that assuming common ground causes speakers to use less words in their narratives than when no common ground can be assumed (because in the first case speakers can rely on their interlocutors to understand implicit references); on the other hand, common ground produces an increase in the use and extent of gestures during speech, possibly to enhance communication with the interlocutors [31, 32, 33]. Finally, gestures may be constrained also by contextual factors, accounting for individual differences, speakers' emotional involvement, etc. These, however, are still largely unexplored.

In L2 communication, L1 gestures appear to have an effect on L2 gestures at all stages of language development. In fact, L2 acquisition is characterized by processes of transfer and interference of gestures from the L1 to the L1 that should be studied, together with verbal language, as part of the interlanguage [35, 36, 37, 38, 39, 40, 41].

Some studies suggest that bilingual speakers might gesture more than monolingual speakers because gesturing helps them formulate their spoken message and is a way to compensate for the reduced proficiency in their L2 [42]. In addition, speakers with low levels of competence might use more L1-specific gestures than speakers with higher levels of competence [40]. L2 speakers' greater use of gestures than L1 speakers might be explained on cognitive grounds, that is, due to the cognitive complexity that speaking a foreign language requires [43].

However, studies do not support unambiguously the idea that bilinguals use more gestures than monolinguals. Other factors besides reduced proficiency in the L2 may account for the differences between the use of gestures in L1 and L2. Communication and contextual factors might affect gesture use in L2 speakers as they do in L1 speakers. For example, common ground might have an effect on L2 speakers' gestures and lead to increased gesturing that is unrelated to L2 speakers' proficiency level [31, 32, 33]. Contextual factors such as task expressiveness, nervousness, as well individual factors might also affect L2 speakers' gestures. Nicoladis et al. [44] examined the relationship between gesture use, L2 proficiency level and task complexity in a story recall task. They found only weak evidence supporting the idea that increased task complexity leads to increased gesture use, and suggest that gesture use might also be related to expressivity, as well to the speaker's gender.

What happens when L2 speakers speak in front of an audience? A number of factors may determine how L2 speakers' use their voice and gestures in a public presentation. Public speaking training classes insist that speakers can improve their non-verbal communication skills by learning the basics and rehearsing before they give their speech in public. It is assumed that rehearsal may help the speaker lessen the tension, sound and look less stiff, more natural during the presentation, and be more pleasant for the listener to hear. For L2 speakers, reducing the tension may significantly impact on the verbal and non-verbal production in L2, and bring about an improvement in both.

There is little scientific research to support the beliefs and assumptions of public-speaking training classes. To fill this gap, this paper reports a preliminary study of students' non-verbal behavior in a presentation in front of a class. The study is part of an investigation aimed at understanding speakers' use of voice and body language in public speaking as well as how non-verbal communication can be enhanced though formal instruction. The study examines the pitch range and gestural characteristics of 3 Italian speakers of English engaged in a story-telling task in Italian and English. The hypothesis tested is that contextual factors such as nervousness or performance anxiety will cause speakers to use narrow pitch range and reduced gesturing; greater ease with the task (because of rehearsal and/or greater familiarity with the task) will cause speakers to use wider pitch range and more gesturing.

3. Experiment

To test the experimental hypothesis, this study compares the pitch variation quotient (PVQ) [2] and the overall number of gestures of three Italian speakers telling the same story, once in Italian and twice in English, as part of an in-class oral presentation task.

3.1. Subjects, Method and Materials

The subjects were part of a larger group of (10) subjects who took part in the experiment. They were all English L2

learners, participating in a public-speaking class, master-degree level, taught by the first author. All subjects were female, mean age 22.75, speakers of Italian L1 and students at the University of Padova, with a competence of English at the B1 level of the CEFR. The data of the remaining 7 subjects are under analysis.

The speakers had to tell the class a fable, Aesop's "The Fox and The Crow", that they had previously read at home. The speakers told the story a first time in Italian, and right afterwards in English. They then repeated the story in English a second time a week later. Thus, the first time the speakers told the fable in Italian and English they had little time to prepare for the task; the second time they had much more time to prepare the story at home before repeating it in class. The speakers were video-recorded by the teacher. Each recording lasted about 90-120 seconds.

The three data sets will be referred to as Italian (=Italian L1); English 1 (=English, repetition at time 1) and English 2 (=English, repetition at time 2).

Out of the whole material, the authors selected 10 utterances that were used by all the subjects telling the fable. In these utterances the concepts expressed were the same, though the words and type of sentences used by the speakers were different. The purpose of selecting only the utterances that were used by all speakers was to compare, for any given utterance, the possible co-occurrence of one or more gesture. The selected utterances are reported in Table 1.

N.	Utterance
1	Once upon a time
2	It was flying around
3	On the shelf of a window
4	It flew down
5	It picked up the cheese
6	It went to the top of the tree
7	The crow opened its beak
8	The cheese fell to the ground
9	The fox caught it
10	It ran away

Table 1: List of utterances selected for the analysis.

3.2. Data Analysis

The audio signal was extracted from the videos using the AVC software (available at http://www.any-video-converter.com/). The audio signal was imported in Praat (www.praat.org), and pitch was measured setting the pitch floor to 75 Hz, and the ceiling to 500 Hz (since all the speakers were female). The boundaries of the selected utterances in the audio files were marked on a text grid. To calculate the PVQ, following a procedure indicated in [2], the pitch listings were extracted from each audio file, the outliers were removed, mean and standard deviation were calculated, and the data were normalized dividing the standard deviation of F_0 by the mean. This procedure was carried out on both the whole audio files and the selected utterances. The statistical significance of the results was tested with one-way ANOVAs with task as a factor, and post-hoc Tukey HSD tests.

The audio signal was then imported in Elan (https://tla.mpi.nl/tools/tla-tools/elan/). An analysis was carried out to annotate each gesture co-occurring with the selected utterances in the three data sets (Italian, English 1 and English 2). At this preliminary stage of analysis, the aim was only to get a total count of the gestures, per speaker and data set, so as to verify if there exists any relation between the variation in the speakers' PVQ and their overall gestures. Because of this, for this analysis, we grouped together all iconic and non-iconic gestures. An analysis of the speakers' gestures classified by type will be carried out in the next phase of the study.

Gesture rate was calculated for each data set following a procedure used in Nicoladis et al. [44]. Gesture rate is a measure of the percentage of word tokens accompanied by gestures, and is calculated by dividing the number of gestures by the total number of words multiplied by a hundred. The use of this measure controls for individual differences in speech.

To calculate the gesture rate for this analysis we counted all the words used in the selected utterances for each speaker. Speakers' disfluencies, repetitions and corrections were computed as part of the total number of words. However, they were also counted separately, as they may reflect grammatical or lexical difficulties that speakers may tend to compensate with their gestures.

4. Results

4.1. Pitch Variation

Tables 2 and 3 show the PVQ data for the three speakers, as calculated, respectively, for the whole story and the selected utterances.

Table 2 shows that all speakers vary their pitch more in the English 2 task than in English 1 or Italian. Interestingly, for all speakers the PVQ of Italian is comparable to the PVQ of English 1, showing that, at time 1, the speakers did not use a very varied pitch in English or Italian. This difference is greater for speaker C than for A or B.

At the ANOVA test, the difference in pitch values in the three tasks was highly significant for all speakers: for speaker A: $F(2, 21421) = 337.06$, $p < .0001$ –though the difference between PVQ in Italian and English 1 was not significant at a Tukey HSD test; for speaker B: $F(2, 17022) = 936.12$, $p < .0001$; for speaker C: $F(2, 24426) = 1724.9$, $p < .0001$.

PVQ - story	Italian	English 1	English 2
Speaker A	0.17	0.18	0.21
Speaker B	0.22	0.23	0.24
Speaker C	0.20	0.20	0.26

Table 2: Pitch variation quotient for the three speakers in the entire story in Italian, English 1 (repetition at time 1) and English 2 (repetition at time 2).

PVQ - utterances	Italian	English 1	English 2
Speaker A	0.18	0.20	0.18
Speaker B	0.22	0.22	0.24
Speaker C	0.23	0.19	0.25

Table 3: Pitch variation quotient for the three speakers in the selected utterances in Italian, English 1 (repetition at time 1) and English 2 (repetition at time 2).

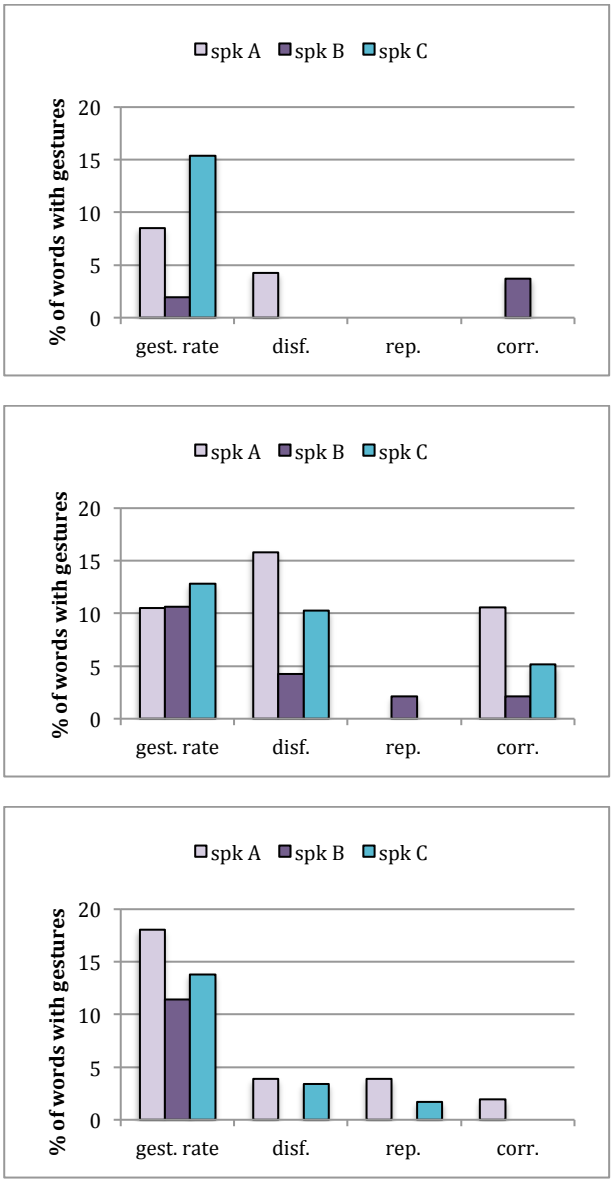
Table 3 shows the PVQ data for the utterances only. Speaker A appears to vary her mean pitch more in English 1 than in the other two data sets, but the difference in PVQ in the three data sets is not significant at the ANOVA test. Speaker B varies her mean pitch more in English 2 than in Italian and English 1 [$F(2,22) = 11.73$, $p = 0.000341$], with a difference between Italian and English 1 that was not significant at the post-hoc Tukey test. Speaker C has higher mean pitch values in Italian

and English 2 than in English 1, but the difference between the three data sets is not significant at the ANOVA test.

4.2. Gesture rate

Figure 1-3 show the gesture rate and percentages of disfluencies, repetitions and corrections for the three speakers in Italian, English 1 and English 2, respectively.

The data show that for two speakers gesture rate increases from Italian to English 1 to English 2; for the third speaker gesture rate is highest in Italian, and then slightly higher in English 2 than in English 1. Disfluencies and corrections are most frequent in English 1, but they occur, for two of the speakers, also in English 2; two speakers show some disfluencies and corrections also in Italian.



Figures 1-3. Gesture Rate, Disfluencies, Repetitions and Corrections in Italian (top), English 1 (center), and English 2 (bottom).

To test the correlation of the present data with the data on the pitch variation we ran Spearman correlation tests, but they did not yield positive correlations, probably because of the limited data provided. However, the data show some trends. Overall, speaker C and A gesture more than speaker B. Speaker C has the highest gesture rate and PVQ in Italian; her gesture rate decreases in English 1 to rise slightly in English 2; her PVQ also decreases in English 1 to rise considerably in English 2. This speaker also has the highest percentage of disfluencies and corrections in the data sets. Speakers A and B show a considerable increase in gesture rate from Italian to English 2. For speaker A, this increase in gesturing cannot be clearly linked to her (non significant) variations in PVQ in the three tasks; however, this speaker shows a high percentage of difluencies, especially in English 1, which might be related to the increase in gesture rate and requires further investigation. Speaker B has the lowest gesture rate in Italian; this rate increases in English 1 and English 2; in English 2 she has shows an increase in PVQ.

5. Discussion and Conclusion

This study is a preliminary investigation of the relationship between speakers' global pitch range and gestures, based on the assumption that their combined effect might contribute to the perception of speakers' liveliness in speech. The data from this study allow us to draw only tentative conclusions, which await confirmation in future studies.

Global pitch range and gesture rate were compared in the speech of 3 native Italian speakers. The speakers told the same story in Italian and in English and then, a week later, in English again. The presentations were part of the students' activities in a public-speaking class.

The analysis shows that when the speakers repeated the story in English the second time their pitch was more varied than when they told the story in Italian and/or English the first time. This is interesting since speakers are expected to show a wider variation in pitch in their native language and not in the L2 –as reviewed in § 1, L2 speech tends to be characterised by limited pitch variation and a narrower pitch range than L1 speech. It is possible that the speakers used a wider pitch range in the second repetition in English due to stylistic and contextual factors. That is, they had more time to prepare, put a greater effort in performing well, had less tension in accomplishing the task, etc. It can be hypothesized that knowing the task, being able to prepare and rehearse for it creates the conditions for sounding more lively in speech. However, we realize that to really evaluate the impact of rehearsal on global pitch range, the experimental design needs to include also a second repetition of the story in Italian. This would allow us to compare the students' performances in the second repetition in Italian and English, and see how pitch range changes with respect to the first repetition in both languages. This will be done in future work.

The gesture data show, as expected, individual differences in the use of gestures. The three speakers show quite different gesture rates in Italian. Also, for speakers A and B gesture rate is lowest in Italian, increases in the first repetition in English, and is highest in the second repetition. For speaker C gesture rate is highest in Italian, it is lowest in the first repetition in English, and rises again in the second repetition in English. Speakers A and B's increased gesture rate in the first repetition in English can be explained on both cognitive and communicative grounds [31, 32, 33]. The speakers may gesture more in English than in Italian because gestures help them tell the story in English L2, which is a complex cognitive activity. At the same time, the speakers may gesture more in English than in Italian because they are adapting their gestures to addressees with whom they share common ground: the speakers are telling the story in front of the class, and the class has heard the story before. Speaker C's lowest gesture rate for

the first repetition in English cannot be attributed simply to cognitive or communicative factors –which would both lead to increased gesturing. Contextual or individual factors, such as the speaker's tension for the task, might have affected her gestures.

Finally, the data show that, in general, speakers' wider pitch co-occur with higher gesture rate, providing preliminary support to our hypothesis.

This study has some obvious limitations, which will be corrected in its continuation. One relevant aspect that this study does not tackle concerns the nature of the gestures produced by the speakers. Future work might show that, for example, L2 speakers produce more deictic gestures in L2 than in L1, as has been shown in much previous research [e.g. 45]. The use of iconic gestures in this task is also worth investigating. Classifying the types of gestures produced by the speakers is indeed important for drawing conclusions in this type of study.

The investigation will be expanded with the addition of more subjects as well as the analysis of other acoustic parameters that might contribute to the perception of speakers' liveliness. Also, the subjects will be tested a second time also in Italian to obtain data that are comparable with second repetition in English.

In spite of its limitations, we believe that this study shows that investigating the relation between global pitch range and gestures in first and second language speech is worth pursuing.

6. References

[1] Hincks, R., "Processing the prosody of oral presentations", Proc. InSTIL/ICALL2004 – NLP and speech technologies in advanced language learning systems, 63-69, 2004.

[2] Hincks, R., "Measuring liveliness in presentation speech", Proc. INTERSPEECH 2005 - Eurospeech, 9th European Conference on Speech Communication and Technology, 01/2005.

[3] Mennen, I., Schaeffler, F. and Docherty, G., A methodological study into the linguistic dimensions of pitch range differences between German and English. Proc. IV Conference on Speech Prosody, University of Campinas, 527-530, 2008.

[4] Mennen, I., Schaeffler, F. and Docherty, G., "Cross-language differences in fundamental frequency range: A comparison of English and German, Journ. of the Acoustical Society of America, 131(3): 2249-2260, 2012.

[5] Graham, C., "Revisiting f0 range production in Japanese-English simultaneous bilinguals", UC Berkeley Phonology Lab Annual Report, 110-125, 2013.

[6] van Bezooijen, R., "Sociocultural aspects of pitch differences between Japanese and Dutch women", Language and Speech, 38(3): 253-256, 1995.

[7] Jenkins, J., "A sociolinguistically-based, empirically-researched pronunciation syllabus for English as an International Language", Applied Linguistics, 23(1): 83-103, 2002.

[8] Aoyama, K. and Guion, S. G., "Prosody in second language acquisition: An acoustic analysis on duration and F_0 range", in O.-S. Bohn and M. J. Munro [Eds], The Role of Language Experience in Second-Language Speech Learning. In Honor of James Emil Flege, 281-297 Amsterdam, John Benjamins, 2007.

[9] Pickering, L., "The structure and function of intonational paragraphs in native and nonnative speaker instructional discourse", English for Specific Purposes, 23: 19-43, 2004.

[10] Traunmüller, H. and Eriksson, A. "The perceptual evaluation of *F0* excursions in speech as evidenced in liveliness estimations", Journ. of the Acoustical Society of America, 97: 1905-1915, 1995.

[11] Ullakonoja, R., "Comparison of pitch range in Finnish (L1) and Russian (L2)", Proc. 16th ICPhS, 1701-1704, 2007.

[12] Johns-Lewis, C., "Prosodic differentiation of discourse modes", in C. Johns-Lewis, [Ed.], Intonation in discourse, 199-220, Breckenham, Kent: Croom Helm, 1986.

[13] Ladd, D. R., Intonational phonology, Cambridge: Cambridge University Press, 1996.

[14] Patterson, D., A linguistic approach to pitch range modeling, PhD dissertation, Univ. of Edinburgh, 2000.

[15] Goldin-Meadow, S., Hearing gesture: How our hands help us think, Cambridge, MA: Harvard University Press, 2003.

[16] Kendon, A., Gesture: visible action as utterance, Cambridge: Cambridge University Press, 2004.

[17] McNeill, D., Hand and mind: What gestures reveal about thought, Chicago: Chicago University Press, 1992.

[18] McNeill, D., Gesture and thought, Chicago: University of Chicago Press, 2005.

[19] Birdwhistell, R. L., Kinesics and context: Essays on body motion communication, Philadelphia: University of Pennsylvania Press, 1970.

[20] Kendon, A., "Gesticulation and speech: Two aspects of the process of utterance", in M. R. Key [Ed.], The Relationship of Verbal and Nonverbal Communication, The Hague: Mouton, 207-227, 1980.

[21] Bull, P. and Connelly, G., "Body movement and emphasis in speech", Journal of Nonverbal Behavior, 9: 169-187, 1985.

[22] Loehr, D. P. Gesture and intonation, Doctoral dissertation, Georgetown University, Dissertation Abstracts International, 65 06, 2180, UMI No. 3137056, 2004.

[23] Esteve-Giberta, N. and Prieto, P., "Prosodic structure shapes the temporal realization of intonation and manual gesture movements", Journal of Speech, Language, and Hearing Research, 56: 850–864, 2013.

[24] Rochet-Capellan, A., Laboissierre, R., Galvan A. and Schwartz, J. "The speech focus position effect on jaw-finger coordination in a pointing task", Journal of Speech, Language, and Hearing Research, 51: 1507-1521, 2008.

[25] de Ruiter, J. P. "The production of gesture and speech", in D. McNeill [Ed.], Language and Gesture, 284-311, Cambridge University Press, Cambridge, 2000.

[26] McClave, E., "Pitch and manual gestures", Journal of Psycholinguistic Research, 27: 69-89, 1998.

[27] Rusiewicz, H. L., "Synchronization of prosodic stress and gesture: A dynamic systems perspective", Proceedings of the 2nd Conference on Gesture and Speech in Interaction (GESPIN 2011), Bielefeld, Germany, 2011.

[28] Roustan, B., and Dohen, M., "Co-production of contrastive prosodic focus and manual gestures: Temporal coordination and effects on the acoustic and articulatory correlates of focus", Proceedings of Speech Prosody 2010, 100110, 1-4, 2010. Retrieved from www.speechprosody2010.illinois.edu/papers/100110.pdf, 2010.

[29] Leonard, T., and Cummins, F., "Temporal alignment of gesture and speech", in Proceedings of the Gesture and

Speech in Interaction (GESPIN 2009) Conference, Poznan, Poland, 2009.
[30] Krahmer, E. and Swerts, M., "The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception", Journal of Memory and Language, 57, 396-414, 2007.
[31] Holler, J. and Wilkin, K., "Communicating common ground: How mutually shared knowledge influences speech and gesture in a narrative task", Language and Cognitive Processes, 24(2): 267-289, 2009.
[32] Holler, J., Tutton, M. and Wilkin, K., "Co-speech gestures in the process of meaning coordination", in Proceedings of the 2nd Conference on Gesture and Speech in Interaction (GESPIN 2011), Bielefeld, Germany, 2011.
[33] Galati, A. and Brennan, S., "Speakers adapt gestures to addressees' knowledge: Implications for models of co-speech gesture", Language, Cognition and Neuroscience, 29(4): 435-451, 2014.
[34] Mol, L., Krahmer, E., Maes, A. and Swerts, M., "Seeing and being seen: The effects on gesture production", Journal of Computer-Mediated Communication, 17(1): 77-100, 2011.
[35] Gullberg, M., "Some reasons for studying gesture and second language acquisition (Hommage à Adam Kendon)", International Review of Applied Linguistics, 44(2): 103-124, 2006.
[36] Pika, S., Nicoladis, E. and Marentette, P. F. "A cross-cultural study on the use of gestures: Evidence for cross-linguistic transfer?", Bilingualism: Language and Cognition, 9: 319-327, 2006.
[37] Brown, A., "Gesture viewpoint in Japanese and English: Cross-linguistic interactions between two languages in one speaker", Gesture, 8(2): 256-276, 2008.
[38] Brown, A. and Gullberg, M., "Bidirectional crosslinguistic influence in L1-L2 encoding of manner in speech and gesture", Studies in Second Language Acquisition, 30(2): 225-251, 2008.
[39] Ortega, L., Understanding second language acquisition, London: Hodder Education, 2009.
[40] Nicoladis, E., "The effect of bilingualism on the use of manual gestures", Applied Psycholinguistics, 28: 441-454, 2007.
[41] Cavicchio, F., and Kita, S., "Bilinguals switch gesture production parameters when they switch languages", Proceedings Tilburg Gesture Research Meeting (TIGeR) 2013, 2013. Retrieved from: http://tiger.uvt.nl/list-of-accepted-papers.html.
[42] Nicoladis, E., Pika, S. and Marentette, P., "Do French-English bilingual children gesture more than monolingual children?", Journal of Psycholinguistic Research, 38 (6): 573-585, 2009.
[43] Kita, S., "How representational gestures help speaking", in D. McNeill [Ed.], Language and gesture, Cambridge: Cambridge University Press, 162-185, 2000.
[44] Nicoladis, E., Pika, S., Yin, H. and Marentette, P., "Gesture use in story recall by Chinese–English bilinguals", Applied Psycholinguistics 28(4): 721-735, 2007.
[45] Sherman, J. and Nicoladis, E., "Gestures by advanced Spanish-English second-language learners", Gesture, 4: 143-156, 2004.