York Papers in Linguistics
Series 2

Issue 11

March 2011

ISSN 1758-0315

Editors:
Lisa Roberts
Mariola Turek
Theodora Lee

Editorial Note

This is the eleventh issue of *York Papers in Linguistics Series 2* (YPL2). It contains papers from staff, graduate students and associated colleagues of the Department of Language and Linguistic Science, University of York.

The editors would like to thank the anonymous reviewers for their invaluable comments and advice during the preparation of this issue, and the editorial board and YPL members for their help and support.

Publication here does not prejudice future appearance elsewhere. For more information on current and past publications of the Department of Language and Linguistic Science, University of York, see http://www.york.ac.uk/language/ypl. All enquiries should be addressed to ypl@york.ac.uk.

Lisa Roberts
Mariola Turek
Theodora Lee

York, 2011

http://www.york.ac.uk/depts/lang/ypl

Contents

COMPARING VOWEL FORMANT NORMALISATION PROCEDURES

NICHOLAS FLYNN

Abstract

This article compares 20 methods of vowel formant normalisation. Procedures were evaluated depending on their effectiveness at neutralising the variation in formant data due to inter-speaker physiological and anatomical differences. This was measured through the assessment of the ability of methods to equalise and align the vowel space areas of different speakers. The equalisation of vowel spaces was quantified through consideration of the SCV of vowel space areas calculated under each method of normalisation, while the alignment of vowel spaces was judged through considering the intersection and overlap of scale-drawn vowel space areas. An extensive dataset was used, consisting of large numbers of tokens from a wide range of vowels from 20 speakers, both male and female, of two different age groups. Normalisation methods were assessed. The results showed that vowel-extrinsic, formant-intrinsic, speaker-intrinsic methods performed the best at equalising and aligning speakers' vowel spaces, while vowel-intrinsic scaling transformations were judged to perform poorly overall at these two tasks.

1. *Introduction*

The process of normalising vowel formant data to permit accurate cross-speaker comparisons of vowel space layout, change and variation, is an issue that has grown in importance in the field of sociolinguistics in recent years. A plethora of different methods and formulae for this purpose have now been proposed.

Thomas & Kendall (2007) provide an online normalisation tool, "NORM", a useful resource for normalising formant data, and one which has opened the viability of normalising to a greater number of researchers. However, there is still a lack of agreement over which available algorithm is the best to use.

This article aims to add to the normalisation literature by comparing a large number of the normalisation procedures available, evaluating their effectiveness at neutralising variation in vowel formant data due to inter-speaker physiological and anatomical differences.

In section 2 the rationale behind normalising vowel formant data is explained, before section 3 presents the formulae of existing methods, and section 4 recounts the findings of other studies that have compared different procedures. Section 5 describes the methodology used for this study and provides information about how the effectiveness of procedures was tested. Section 6 gives the results of the comparisons between the different methods, before section 7 discusses the relative effectiveness of each procedure and compares the findings to those of previous comparative studies of normalisation methods.

Throughout this article the following notation is used: F_i represents a formant, (i = 1,2,3). A superscript ^{N} is used to denote a normalised value. For example, F_1^N is the normalised value of the first formant of a token.

2. *The concept of vowel formant normalisation*

A major problem faced by researchers in sociophonetic variation is that no two speakers' vowel tracts share the same dimensions. As a consequence, the "same" phonological vowel uttered by different speakers will show formants at different frequencies due to the different sizes of the speakers' vocal tracts. For example, female speakers tend to display higher formant frequencies than male speakers, as their vocal tracts are shorter and thus their resonance frequencies are higher. It can be difficult, then, when comparing the positioning of vowels within speakers' vowel spaces, to identify whether differences in formant values are due to a linguistic change in the vowel system, or are merely due to the anatomical and physiological differences between speakers.

It has been acknowledged that the raw Hertz formant frequencies of different speakers are not directly comparable, and that it is not ideal to plot formant values in Hertz from different speakers on the same formant chart (Watt et al. 2010). This presents a problem for sociophonetic research that seeks to describe variation and change through the comparison of speech from different speakers.

The solution is, in principle, to remove as much of the inter-speaker formant value differences due to biological differences as possible. This would leave quantities unaffected by the size of a speaker's vocal tract, and so would be directly comparable. The process of transforming formant frequencies to make them directly comparable with those from other speakers is called Vowel Formant Normalisation.

A number of differing formulae have been put forward as normalising algorithms. The sheer number of proposed normalisation algorithms indicates a distinct lack of consensus about how best to normalise. However, those researchers who have considered the process of normalisation have collectively identified a number of goals of normalisation:

1. to minimise or eliminate inter-speaker variation due to inherent physiological or anatomical differences;
2. to preserve inter-speaker variation due to social category differences, including age, gender and dialect, or due to sound change;
3. to maintain vowel category and phonemic differences;
4. to model the cognitive processes that allow human listeners to normalise vowels uttered by different speakers.

(Hindle 1978; Disner 1980; Thomas 2002; Langstrof 2006; Thomas & Kendall 2007; Fabricius 2008; Clopper 2009; Watt et al. 2010).

Of course, such goals are somewhat idealistic, and the unlikelihood of any normalisation method perfectly fulfilling all the above criteria has been acknowledged (Thomas 2002; Adank et al. 2004; Bigham 2008; Thomas & Kendall 2007). For example, it has been observed that through reducing physiological variation, sociolinguistic variation can also be reduced (Adank et al. 2004).

Some researchers may place greater importance on one criterion over the others. This will largely depend on the nature of the study. For example, perception-based studies want normalisation to approximate the process of human vowel perception as closely as possible (Rosner & Pickering 1994; Syrdal & Gopal 1986), while sociophonetic studies are less concerned with this, but place greater importance on the maintenance of sociophonetically-relevant information, such as age-based variation (Fabricius 2008; Watt et al. 2010; Langstrof 2006; Thomas & Kendall 2007).

Thomas (2002) makes the excellent observation that

> all normalisation techniques have drawbacks, […] choosing which normalisation technique to use is a matter of deciding which drawbacks are tolerable for the study at hand.

Thomas (2002:174)

The onus, then, appears to be on the researcher to choose from the numerous posited methods, a normalisation procedure that is appropriate for the type of study and its research objectives.

3. *Existing normalisation formulae*

Normalisation procedures have traditionally been categorised according to whether they are vowel intrinsic or extrinsic, formant intrinsic or extrinsic, speaker intrinsic or extrinsic, or a combination of these six categories (Adank 2003; Adank et al. 2004; Thomas & Kendall 2007; Clopper 2009; Fabricius et al. 2009; Watt et al. 2010).

Vowel-intrinsic techniques use information from a single vowel token, while vowel-extrinsic techniques use information from multiple vowels, often across several vowel categories to normalise a formant value. Formant-intrinsic procedures normalise a formant value using information from occurrences of that formant only, using F_1 measurements to normalise an F_1 value, for example, while formant-extrinsic procedures use information from multiple formants, for example using F_1, F_2 and F_3 measurements to normalise an F_1 value. Speaker-intrinsic methods use information from a single speaker, while speaker-extrinsic methods use information from a population.

Speaker-extrinsic procedures are rarely used, due to their complexity, and the fact that by their very nature, adding more speakers into a dataset will alter the normalised values, meaning that any calculations that have already been made must be discounted and redone from scratch. Despite this disadvantage, Labov et al. (2006), a recent major piece of American English sociolinguistic work, used a speaker-extrinsic normalisation procedure as part of its methodology.

A number of speaker-intrinsic normalisation procedures will now be presented, beginning with vowel-intrinsic methods. These can be regarded as a rescaling of Hertz frequencies onto a different scale, and were originally developed to closer model human vowel perception by transforming frequencies onto a more perceptually-relevant scale (Adank 2003; Adank et al. 2004).

A number of scales have been proposed, including the Mel scale, obtained via (1) (Stevens & Volkmann 1940), Equivalent Rectangular Bandwidth (ERB), obtained via (2) (Glasberg & Moore 1990) and Bark, obtained using (3) (Traunmüller 1990).

(1) $$F_i^N = 1127 \ln.\left(1 + \frac{F_i}{700}\right)$$

(2) $$F_i^N = 21.4 \ln.(0.00437F_i + 1)$$

(3) $$F_i^N = 26.81 \left(\frac{F_i}{1960 + F_i}\right) - 0.53$$

Bladon et al. (1984) extended the Bark transformation to normalise female speakers relative to male speakers using (4). Clopper (2009) makes the justifiable criticism that subtracting an extra 1 Bark for female speakers appears an arbitrary choice.

(4)
$$F_i^N = \begin{cases} 26.81\left(\dfrac{F_i}{1960 + F_i}\right) - 0.53 & \text{, speaker is male} \\ \left(26.81\left(\dfrac{F_i}{1960 + F_i}\right) - 0.53\right) - 1 & \text{, speaker is female} \end{cases}$$

Syrdal & Gopal (1986) used the Bark transformation to create a formant-extrinsic, vowel-intrinsic method called the "Bark-Distance Method". Their method is based on their observation that the distance between neighbouring formants is similar across speakers (Syrdal & Gopal 1986). The equations used to normalise using Syrdal & Gopal's Bark-Distance Method are:

(5) $F_1^N = F_1^{BARK} - F_0^{BARK}$

(6) $F_2^N = F_3^{BARK} - F_2^{BARK}$

The correlates of vowel height and vowel frontness for F_1 and F_2 are maintained using these measurements (Syrdal & Gopal 1986).

Miller (1989) proposed scaling formant frequencies to a scale better aligned with perceptual differences by taking the (natural) logarithm of the Hertz value. That is,

(7) $F_i^N = \ln.(F_i)$

He then extended this concept, by suggesting a formant ratio model, broadly given in (8). *SR* is defined as the "Sensory Reference", a speaker-specific value based on the average fundamental frequency of the speaker whose frequencies are being normalised, and of all speakers in the sample (Miller 1989).

(8)
$$F_i^N = \begin{cases} \left(\dfrac{\ln.(F_i)}{\ln.(F_{i-1})}\right) & ,\ i > 1 \\ \left(\dfrac{\ln.(F_i)}{SR}\right) & ,\ i = 1 \end{cases}$$

Miller's formant ratio method is an example of a speaker-extrinsic, formant-extrinsic, vowel-extrinsic procedure, as *SR* is derived from measurements taken from a population of speakers. Another speaker-extrinsic, formant-extrinsic, vowel-extrinsic procedure has been proposed by Nordström (1977).

Nordström's method uses (9) to scale formant values based on using an estimation of the difference between male and female vocal tract length to transform female speakers' values relative to males'.

(9)
$$F_i^N = \begin{cases} F_i & \text{, speaker is male} \\ \left(\dfrac{\mu_{F_3}^{male}}{\mu_{F_3}^{female}}\right) F_i & \text{, speaker is female} \end{cases}$$

Here, $\mu_{F_3}^{male}$ and $\mu_{F_3}^{female}$ are the mean F_3 of all tokens with $F_1 > 600$Hz for all male speakers and all female speakers in the sample respectively.

Many formulae exist that are speaker-intrinsic, formant-intrinsic, and vowel-extrinsic. An early example which is recounted in Lobanov (1971), was the linear compression and expansion method, which scales a speaker's formant values relative to their maximum value for that formant.

(10) $$F_i^N = \frac{F_i}{F_i^{max}}$$

This allows comparability between speakers, as formant values are represented as a proportion of a speaker's maximum formant frequency. In effect, speakers' vowel spaces are aligned by anchoring them at the maximum values for individual formants. Gerstman's (1968) method builds on this methodology by aligning speakers' vowel spaces at both endpoints of their formant frequency range. Values are scaled so that the values of the extremities are 0 and 999 rather than 0 and 1. The equation for Gerstman's method is given in (11).

(11) $$F_i^N = 999\left(\frac{F_i - F_i^{min}}{F_i^{max} - F_i^{min}}\right)$$

The maximum and minimum F_i in (10) and (11) are taken from all vowel tokens for a speaker.

The previous two normalisation procedures express values relative to the extremities of a speaker's formant frequency range. The next procedure, developed by Watt & Fabricius (2002) expresses values relative to the constructed centroid of a speaker's vowel space.

In this technique, a speaker's vowel space is thought of as a triangle, with the apices at points representing the minimum and maximum F_1 and F_2 for the speaker, and labelled [i], [a] and [u′]. The coordinates of [i] are the speaker's mean F_1 and F_2 for vowels taken from their FLEECE lexical set, which are used to represent their minimum F_1 and maximum F_2 respectively. The coordinates of [a] are the speaker's mean F_1 and F_2 for their TRAP lexical set, with the F_1 value here representing the speaker's maximum observed F_1. [u′] is constructed so that $F_1[u'] = F_2[u'] = F_1[i]$, and represents the minimum F_1 and F_2 possible for a speaker.

The centroid, *S*, of this triangle is then found as the grand mean of points [i], [a] and [u′] using (12). Figure 1 illustrates this.

(12) $$S(F_i) = \frac{F_i[i] + F_i[a] + F_i[u']}{3}$$

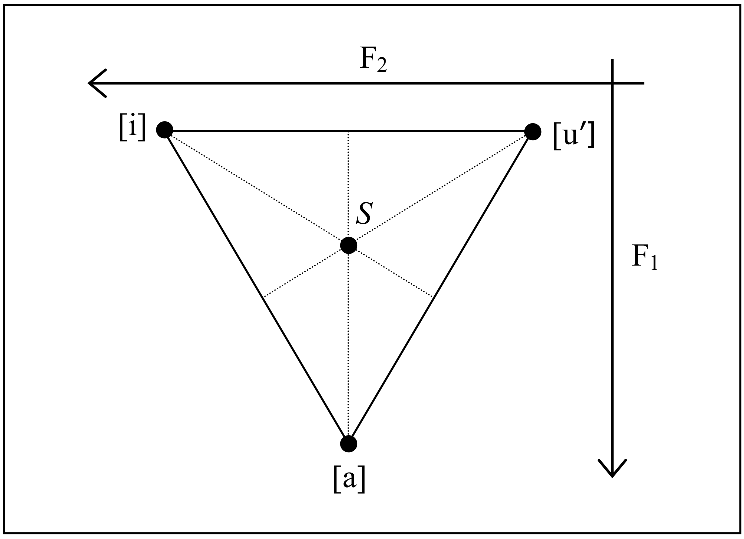


Figure 1: Construction of the centroid S as part of the Watt & Fabricius method of normalisation

Formant values are then expressed relative to the centroid.

(13) $F_i^N = \dfrac{F_i}{S(F_i)}$

It has been acknowledged that this procedure can skew values in the lower part of the vowel space (Thomas & Kendall 2007; Fabricius et al. 2009; Bigham 2008). As a result, Fabricius et al. (2009) offer a modified formula for calculation of the coordinates of the centroid.

(14) $S(F_i) = \begin{cases} \dfrac{F_i[\text{i}] + F_i[\text{a}] + F_i[\text{u}']}{3} & , i = 1 \\ \dfrac{F_i[\text{i}] + F_i[\text{u}']}{2} & , i = 2 \end{cases}$

Other variations of the Watt & Fabricius method have been used in research projects. For example, Kamata (2008) used mean formant values of the KIT and START lexical sets rather than the FLEECE and TRAP sets to construct the apices of the speakers' vowel triangles, because in the variety she was studying, FLEECE was subject to diphthongisation, and TRAP was suspected to be undergoing a shift (Kamata 2008).

Bigham (2008) used the centroid of a quadrilateral rather than a triangle for his research, as he believed a quadrilateral shape was a better reflection of the vowel space of American English than a triangle (Bigham 2008). The four apices of the quadrilateral used were the mean formant values for a speaker of the American English vowels [ɪ], [u], [æ] and the average of [ɑ] and [ɔ], with tokens taken from word list items of the form /h*V*d/. As per the Watt & Fabricius method, to normalise, a speaker's formant values were expressed relative to their respective centroid, using (13).

A further normalisation procedure that expresses values relative to the hypothetical centre of a speaker's vowel space is that developed by Lobanov (1971). Using a method similar to that in statistics to transform normally distributed data to a uniform normal distribution, formant values are normalised by subtracting a speaker's mean formant value across all vowel tokens, and then dividing by the standard deviation[1] for the formant across all vowels for that speaker.

(15) $$F_i^N = \frac{(F_i - \mu_i)}{\sigma_i}$$

The final technique to be presented is attributed to Nearey (1978). It has two formulations, one formant-intrinsic and the other formant-extrinsic. In both cases, talkers' vowel spaces are made comparable by aligning them at speakers' mean formant frequencies (Clopper 2009).

In the formant-intrinsic formulation, sometimes referred to as Nearey's Single Log-Mean Method (Adank et al. 2004) or Nearey's Individual Formant Mean Method (Clopper 2009), the natural logarithm of a speaker's formant value is taken, and then the mean of the log-transformed formant frequency across all vowels for the speaker is subtracted.

(16) $$F_i^N = \ln.(F_i) - \mu_{\ln.(F_i)}$$

In the formant-extrinsic formulation, sometimes called Nearey's Grand-Mean Method (Clopper 2009) or Nearey's Shared Log-Mean Method (Adank et al. 2004), the natural logarithm of a speaker's formant value is taken, and then the mean of the log-transformed formant frequency of all formants of all vowels for the speakers is subtracted.

(17) $$F_i^N = \ln.(F_i) - \mu_{\ln.(F_j)} \quad , \ \forall\, j = 1, \ldots, n$$

The results of studies which have compared the outcomes of normalising via different methods will now be considered.

4. Findings of previous comparative studies

Most researchers who devised their own normalisation methods, compared their normalised results to raw Hertz formant values, and often, values resulting from a transformation of the Hertz values onto a different scale, to evidence the improvement normalising via their formula offered in making formant values from different speakers comparable and giving weight to their argument that adoption of their normalisation formula is warranted.

For example, Watt & Fabricius (2002) showed their method dramatically improved the area ratio and degree of overlap of vowel spaces from two different speakers in comparison to raw Hertz values and Bark-transformed values.

Lobanov (1971) compared normalising using his formula to using Gerstman's method and to using the linear compression and expansion technique. He found that his own method performed the best of the three at reducing the spread of points of the same vowel spoken by different speakers while at the same time maximising the distances between adjacent phonemically-opposing vowels.

Fabricius et al. (2009) compared Watt & Fabricius' method with Nearey's individual log-mean method and Lobanov's method using very rigorous methodology and statistical testing of improvements of Hertz data and of differences between methods in their success at normalising. Their conclusion, was that the Watt & Fabricius method performed at least as

[1] In the original formulation, Lobanov (1971), RMS deviation rather than standard deviation is used.

well as Lobanov's and Nearey's methods, although the results actually show that Lobanov's method outperformed the other two methods overall to a statistically significant extent.

There is always the danger, when researchers compare their own methodology to that of others, that the experiment or results will be somewhat biased towards showing their own procedure in the best light. However, a number of independent comparisons of different normalisation methods have also been conducted.

One of the earliest was Hindle (1978), who compared normalising by Nordström's method, by Nearey's individual log-mean method, and by a six parameter regression method (see Hindle 1978:166 for details). Hindle (1978) concluded that Nearey's method performed the best of the three at normalising data overall, despite it not clustering formant values for the same vowel from different speakers as closely as the six parameter regression method, because Nearey's method was most successful at preserving known age-related formant differences.

Disner (1980) also found Nearey's method to be the best normalisation technique for English vowels out of the four she evaluated, although she used Nearey's grand-mean method.

More recently, Clopper (2009) evaluated both Nearey's individual log-mean and grand-mean methods along with seven other procedures, and found the two to be equally good at producing highly overlapping vowel spaces and aligning the vowel categories of different speakers. Clopper (2009) also found Lobanov's, Watt & Fabricius' and Gerstman's methods to be successful and effective techniques. Rather than singling out one specific procedure as performing the best, she concluded that it is vowel-extrinsic methods in general that are most effective, an opinion that is corroborated by Fabricius et al. (2009) and was earlier posited by Adank (2003) and Adank et al. (2004). Adank et al. (2004) further claimed that specifically, formant-intrinsic rather than extrinsic, vowel-extrinsic methods are the best to use for language variation research. They based this proposal on the findings of their comparison of seven procedures as well as four vowel-intrinsic scaling formulae.

Unlike Clopper (2009), Adank et al. (2004) did indicate a method that performed the best overall, namely, Lobanov's procedure, which they found to be most efficient at preserving phonemic variation, and joint best with Nearey's individual log-mean method at removing variation due to physiological differences (Adank et al. 2004). Langstrof (2006) also came to the conclusion that Lobanov's method was the best technique.

The evidence appears to be, then, that Lobanov's method and one or other of Nearey's methods are consistently found to be the most effective procedures when it comes to normalising vowel formant data.

In coming to their conclusions of which normalisation method is the best to use, studies varied as to how exactly a procedure was evaluated. Some studies looked for overall improvement of vowel space overlap, while some looked for a closer clustering of different speakers' realisations of the same vowels. Also, the purpose of normalising data varied, with the earlier studies such as Disner (1980) and Hindle (1978) seemingly angled more towards normalising to replicate human vowel perception rather than to aid presentation of vowel formant data in sociolinguistic research.

In addition, the range of techniques tested in each study was variable. Hindle (1978) did not test Lobanov's method, while Adank et al. (2004) omitted Watt & Fabricius' method from their otherwise notably thorough investigation. Langstrof (2006) and Fabricius et al. (2009) each only considered three procedures altogether. Clopper (2009) compared an impressive range of methods, but only used data from two speakers. Moreover, she came to her conclusions without any statistical testing or rigorous mathematical procedure. For these reasons, the robustness of her results could be called into question.

A final point to make, is that very few of the existing comparative studies utilised British English data, the exception being Fabricius et al. (2009). Adank (2003) and Adank et al. (2004) used Dutch data, while Lobanov (1971) used Russian. Indeed, Lobanov's method was originally formulated to aid classification of Russian vowels.

Disner (1980) normalised data from six different Germanic languages by a variety of methods, and found that no one single procedure could, in her opinion, be considered the best at normalising for all the languages tested. For this reason, it could be argued that the normalisation procedure found most effective at normalising Dutch formant values by Adank et al. (2004), need not be the most effective procedure for normalising English data. Furthermore, the most effective procedure for normalising American English data need not be the most effective at normalising British English data (and vice versa).

The remainder of this article presents the methodology and results of a comparative study that used a sizeable dataset of British English data to compare a large variety of differing normalisation techniques, both older techniques that some newer comparative studies did not include due to the assumption that they are outdated, and more recently-proposed methods.

5. Methodology

The data that were normalised came from 20 speakers, 5 young (aged 18-22) and 5 older (aged 40-50) speakers of each sex, all resident in Nottingham. F_1 and F_2 measurements were extracted from monophthongal word list items, taken from the word list tasks of longer one-to-one sociolinguistic interviews collected for Flynn (fc). Mono tracks sampled at 22,050Hz were used, and formant measurements were taken in Praat using a Praat script[2]. A minimum of three adjacent points plotted by Praat's inbuilt formant-tracking tool were averaged for each formant of the vowel. Formant values were measured in Hertz, and then normalised using the algorithm of each procedure under comparison. As not all the procedures under consideration are available as part of NORM, the decision was made to use a specially-prepared Microsoft Excel spreadsheet to perform all the normalisation calculations. 3605 tokens were normalised altogether, giving an average of 180 tokens per speaker. Following normalisation, individual tokens were categorised according to their vowel keyword category, (Wells 1982), and a mean normalised value for each keyword category was calculated for each speaker.

20 procedures of vowel formant normalisation were compared, of which 6 were vowel-intrinsic scaling transformations, and the remaining 14 were vowel-extrinsic. Table 1 summarises the methods used. The equations for all 20 methods can be found in the appendix. Many of the procedures were introduced in section 3. Where modifications were made to established methods, these are described below, as are additional procedures not defined in section 3.

Two logarithmic transformations were computed. One was performed with base 10, the other was the natural logarithm function. Following the method of recent work by Adank (2003), Adank et al. (2004) and Clopper (2009), the equation given by Traunmüller (1990) was used to transform data from Hertz to Barks for all procedures using the Bark scale. A Bark-Difference measure was completed, following the idea of Syrdal & Gopal (1986), but using

[2] The Praat script used was devised by Phil Harrison, originally for use in forensic casework.

Procedure	**Hereafter referred to as**	**Topographical Classification**		
		Formant	**Vowel**	**Speaker**
Log transformation (uses log to the base 10)	*Log*	Intrinsic	Intrinsic	Intrinsic
Ln transformation (uses the natural logarithm)	*Ln*	Intrinsic	Intrinsic	Intrinsic
ERB transformation	*ERB*	Intrinsic	Intrinsic	Intrinsic
Mel scale transformation	*Mel*	Intrinsic	Intrinsic	Intrinsic
Bark scale transformation	*Bark*	Intrinsic	Intrinsic	Intrinsic
Bladon et al. method	*Bladon*	Intrinsic	Intrinsic	Intrinsic
Bark-difference method	*Bark-diff*	Extrinsic	Intrinsic	Intrinsic
Linear Compression/Expansion method	*LCE*	Intrinsic	Extrinsic	Intrinsic
Gerstman method	*Gerstman*	Intrinsic	Extrinsic	Intrinsic
Lobanov method	*Lobanov*	Intrinsic	Extrinsic	Intrinsic
Nordström method	*Nordström*	Extrinsic	Extrinsic	Extrinsic
original Watt & Fabricius method	*origW&F*	Intrinsic	Extrinsic	Intrinsic
Watt & Fabricius method modified as in Fabricius et al. (2009)	*1mW&F*	Intrinsic	Extrinsic	Intrinsic
Watt & Fabricius method modified as described below	*2mW&F*	Intrinsic	Extrinsic	Intrinsic
modified version of Bigham's method	*Bigham*	Intrinsic	Extrinsic	Intrinsic
lettER method	*Letter*	Intrinsic	Extrinsic	Intrinsic
Nearey's individual log-mean method	*Nearey1*	Intrinsic	Extrinsic	Intrinsic
Nearey's grand-mean method	*NeareyGM*	Extrinsic	Extrinsic	Intrinsic
Nearey's individual log-mean method as implemented in NORM	*exp{Nearey1}*	Intrinsic	Extrinsic	Intrinsic
Nearey's grand-mean method as implemented in NORM	*exp{NGM}*	Extrinsic	Extrinsic	Intrinsic

Table 1: The normalisation procedures included in this comparative study

the modification as applied by Thomas & Kendall (2007), namely, that $B_3 - B_1$ is substituted in place of $B_1 - B_0$ in Syrdal & Gopal's (1986) original methodology. (B_i represents Bark-transformed F_i.) The formula for computing the normalised values via this procedure can be expressed as (18).

(18) $F_i^N = B_3 - B_i \quad , \quad i < 3$

Nordström was implemented through the use of (9), with $\mu_{F_3}^{male}$ and $\mu_{F_3}^{female}$ defined as the mean F_3 calculated across all vowel tokens having $F_1 > 600$Hz from all male speakers and from all female speakers respectively. For the dataset used, $\mu_{F_3}^{male}$ was found to be 2494Hz, and $\mu_{F_3}^{female}$ was found to be 2847Hz.

The scale factor for female speakers, $\frac{\mu_{F_3}^{male}}{\mu_{F_3}^{female}}$ was therefore calculated to be 0.876 (to 3dp). [3]

As can be seen in Table 1, the Watt & Fabricius method was implemented in three different formulations. In each case, the general formula given in (19) was used to normalise the formant values.

(19) $F_i^N = \frac{F_i}{S(F_i)}$

S is defined as the centroid of a triangle with the apices of the triangle denoted [i], [a] and [u′] derived from the raw Hertz formant values of a speaker. (See Figure 1 in section 3 for an illustration of this methodology.) The construction of [i], [a] and [u′] differed for each of the three techniques.

OrigW&F used the original technique from Watt & Fabricius (2002). Under this method, F_i[i] = F_i[FLEECE], and F_i[a] = F_i[TRAP], using the mean points of the respective keywords. [u′] is then constructed so that

(20) $F_1[u'] = F_2[u'] = F_1[i]$

and the formant values of the centroid, *S*, are calculated using (21).

(21) $S(F_i) = \frac{F_i[i] + F_i[a] + F_i[u']}{3}$

1mW&F implemented the minor modification to the calculation of *S* introduced by Fabricius et al. (2009) in response to Thomas & Kendall's (2007) comment that the original Watt & Fabricius formula can distort the lower part of the vowel space. This adjustment places *S* equidistant between [i] and [u′] on the F_2 axis. The formula used to derive the formant values of *S* using this modified method is given in (22). Following Thomas & Kendall (2007), the following additional modifications were made to Watt & Fabricius' original formula in arriving at *1mW&F*. Firstly, rather than using F_1 and F_2 of mean FLEECE, F_1[i] was set equal to the F_1 of whichever mean keyword vowel had lowest F_1, and F_2[i] was set equal to the highest F_2 value of the mean keyword vowels. Similarly, F_1[a] was taken from whichever keyword vowel category had the highest mean F_1. F_2[a] was not computed, as it isn't used in the calculation of $S(F_2)$ under this methodology. [u′] was constructed using (20), as before.

[3] The scale factor used in the normalisation calculations was not rounded.

(22) $$S(F_i) = \begin{cases} \dfrac{F_i[i] + F_i[a] + F_i[u']}{3}, & i = 1 \\ \dfrac{F_i[i] + F_i[u']}{2}, & i = 2 \end{cases}$$

For *2mW&F*, [i] and [a] were constructed following identical procedure to that for *1mW&F*. However, [u′], rather than having formant values derived from the point [i], was constructed such that F_2[u′] was set equal to the lowest F_2 value of the mean keyword vowels, and F_1[u′] was set equal to the lowest F_1 value of the mean keyword vowels. The decision to derive [u′] in this way was made because it gives an arguably more realistic placement of [u′] in a speaker's vowel space. Using (20) results in [u′] having an F_2 value far lower than a vowel ever would have in reality. However, based on the results of Watt & Fabricius (2002), Fabricius et al. (2009) and Clopper (2009), deriving [u′] in this way appears to result in well-normalised data. Inclusion of *2mW&F*, with [u′] constructed as described, was intended to see whether this improved, worsened or had little or no effect on normalised data than when [u′] is constructed using (20).

A further derivation of the original Watt & Fabricius method that was included, was a modified version of Bigham's (2008) method. As per the three versions of the Watt & Fabricius method, a speaker's formant values were normalised using (19). However, under this method, *S* was equated from 4 points, that is, *S* is the centroid of a quadrilateral rather than a triangle. This method was included to see the effects on normalised data when *S* is derived from four points rather than three. For ease of notation, the points shall be denoted as [i′], [a′], [o′] and [u′].

The formant values of these constructed points were based on a speaker's mean formant values of the keyword vowel categories. F_2[i′] was set equal to the maximum F_2 value, while F_1[i′] was set equal to the minimum F_1 value, as was F_1[u′]. F_2[u′] was set equal to the minimum F_2 value, as was F_2[o′], while F_1[o′] and F_1[a′] were both set equal to the maximum F_1 value. F_2[a′] was set equal to F_2[TRAP].

The formant values of *S* were then calculated using (23), an equation equivalent to, but much simpler than, those suggested by Bigham (2008).

(23) $$S(F_i) = \frac{F_i[i'] + F_i[a'] + F_i[o'] + F_i[u']}{4}$$

It could be argued that a quadrilateral shape better reflects the vowel space of British English speakers than a triangle, as speakers will have both front and back maximally-low vowels (for example TRAP and START respectively) rather than one central maximally-low vowel which a triangular representation of the vowel space could be viewed as implying.

A further innovative technique that was included was *Letter*. Like the Watt & Fabricius method and its derivations, this method fixes a speaker's vowel space at its midpoint and then expresses formant values in relation to this point. However, rather than constructing the formant values of the midpoint as the Watt & Fabricius method and derivations do, actual mean formant values of a speaker's *lett*ER vowel are used.

This is justifiable, as *lett*ER is typically realised as an unstressed schwa-like vowel in English (Wells 1982) and schwa has been described as defining the midpoint of a speaker's vowel space (Johnson 2003; IPA 1999).

It should be noted, that *lett*ER is not always realised as schwa. For example, a lowered [ɐ]-like pronunciation has been reported for speakers from Tyneside (Watt & Milroy 1999; Watt & Allen 2003; Beal 2008), London (Tollfree 1999; Trudgill 1986), and Birmingham and the Black Country (Clark 2008), while a lowered and retracted [ɒ]-like pronunciation has been recounted as appearing in the speech of speakers from Sheffield (Beal 2008), Manchester (Beal 2008) and Nottingham (Flynn 2007). Therefore, caution should be taken when normalising via this method, especially when speakers speak a variety where *lett*ER is not always schwa-like. As a result, care was taken to only include tokens perceived auditorily as [ə] in the calculations of the mean F_1 and F_2 of a speaker's *lett*ER vowel as used in this study.

To normalise the data relative to *lett*ER, the formula given in (24) was used.

(24) $$F_i^N = \frac{F_i}{F_i[lett\text{ER}]} - 1$$

The similarity to (19) is apparent, with the added action of subtracting 1 from the resulting ratio of the raw formant value to corresponding *lett*ER formant. This serves to centre the normalised vowel space at (0,0) and to allow the relative positioning of vowels in relation to the vowel space midpoint to be more easily interpreted in acoustic-phonetic terms from the normalised values. Specifically, a negative F_1 indicates a position higher than centre, while a positive F_1 indicates a position in the vowel space lower than centre. Also, a positive F_2 value represents a fronter than central position in the vowel space, while a negative F_2 value is indicative of a backer than centre position.

The final methods that were included were algorithms based on Nearey (1978). It was noticed that when NORM normalises via either of Nearey's (1978) methods, it further takes the exponential[4] of each result calculated through the use of the original formulae. It could be argued that such an action might reverse some of the effects of the normalisation, since the exponential function is the inverse function of the natural logarithm. Certainly, additionally taking the exponential expresses values on a different scale.

The decision was therefore made to calculate the exponential of Nearey-normalised values to create two further sets of normalised data to compare with the other results, and in particular, to those of Nearey's (1978) original formulae, to see if taking the exponential improved, worsened or had no effect on the effectiveness of normalising via one of Nearey's original methods. The equations used can be expressed as (25) for *exp{Nearey1}*, which has *Nearey1* as a basis, and (26) for *exp{NGM}*, which has *NeareyGM* method as a basis.

(25) $$F_i^N = \exp.\{\ln.(F_i) - \mu_{\ln.(F_i)}\}$$

(26) $$F_i^N = \exp.\left\{\ln.(F_i) - \left(\frac{\mu_{\ln.(F_1)} + \mu_{\ln.(F_2)}}{2}\right)\right\}$$

To reveal the overall effectiveness of each procedure, a series of comparisons were performed to test the efficiency of techniques at satisfying the goals of normalisation. In this article, results are reported for the relative ability of each method in equalising and aligning the vowel space areas of the 20 speakers.

Equalisation of different speakers' vowel spaces could be seen as indicating the removal of variation attributable to anatomical differences between speakers, leaving directly

[4] Thomas & Kendall (2007) use the equivalent term 'anti-log' rather than 'exponential'.

comparable vowel spaces of similar dimensions. Taking inspiration from Fabricius et al. (2009), the equalisation of vowel space areas was quantified by examining the reduction of variance in the speakers' vowel space areas as calculated under each normalisation procedure. A speaker's vowel space was taken to be quadrilateral, with the apices constructed from mean keyword category formant values calculated under each method for each speaker, in a similar way to quadrilaterals involved in *Bigham*. That is, four points were defined, one with values (max. F_2, min. F_1), one (min. F_2, min. F_1), one (min. F_2, max. F_1) and one (F_2[TRAP], max. F_1). The lines connecting these four points were taken to represent the hypothetical outer limits of a speaker's vowel space. A graphical depiction of this is shown in Figure 2.

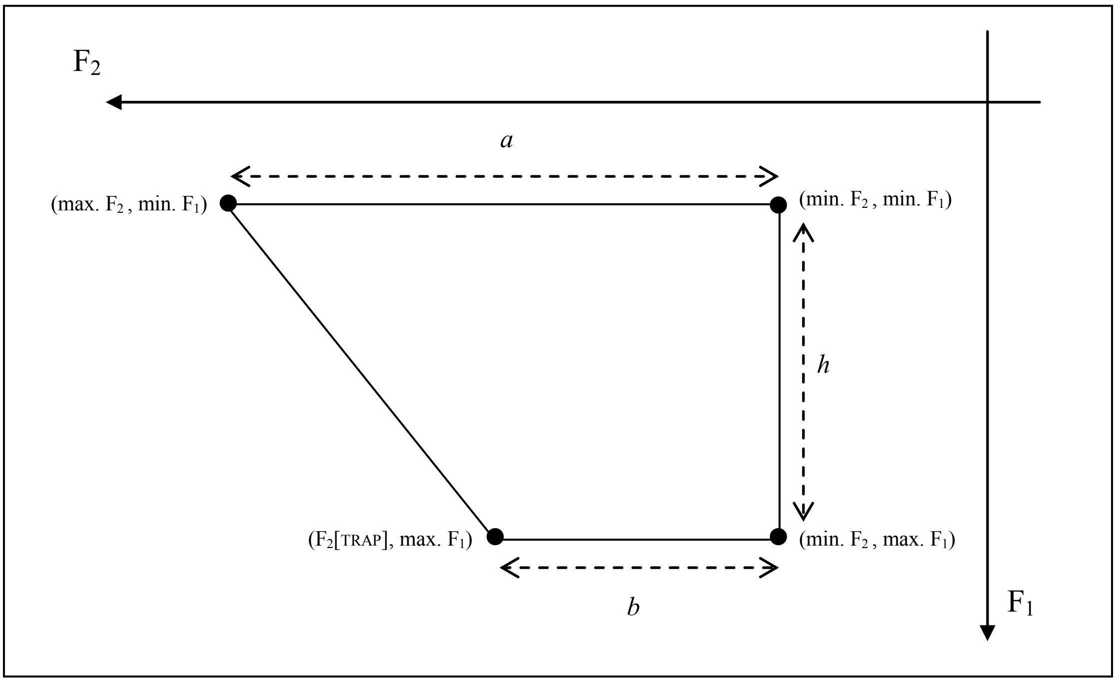


Figure 2: Construction of quadrilateral speaker vowel space areas

The generic formula for calculating the area of a trapezium given in (27) was used to calculate the area of each vowel space.

(27) $A = ½(a+b)h$

The vertical height, h, as shown in Figure 2, was taken to be the F_1 range, while the parallel sides, a and b, were taken to be the F_2 range at the minimal and maximal extremities of the F_1 range respectively. Equations (28) to (30) present numerically the calculation of the dimensions used in the application of (27).

(28) $a = \text{max. } F_2 - \text{min. } F_2$

(29) $b = F_2(\text{TRAP}) - \text{min. } F_2$

(30) $h = \text{max. } F_1 - \text{min. } F_1$

As the different normalisation methods give normalised values in different units, the resulting vowel space areas also differed in terms of units. It was therefore not possible to immediately compare directly the amount of variance existing in the sets of vowel space areas under different normalisation procedures to see which method best minimised variance and thus equalised areas to the greatest extent. Following the method of Fabricius et al. (2009), the squared coefficient of variance (SCV), a scale-invariant measure, was calculated for each procedure using (31). The SCVs were then compared.

(31) $$SCV = \left(\frac{\sigma}{\mu}\right)^2$$

A low SCV is indicative of a dataset having small variance, while a high SCV indicates a dataset has large variance. It is simple then to rank procedures in their effectiveness at equalising areas through examining the respective SCVs. Any normalisation method resulting in areas with a lower SCV than that of the SCV of the raw Hertz areas can be said to have reduced the variance of inter-speaker vowel space areas, and hence made different speakers' vowel space areas more similar. Furthermore, the method resulting in areas with the lowest SCV overall can be said to have reduced area variance the most, and hence equalised the vowel space areas of different speakers to the greatest extent.

In addition to considering the equalisation of vowel space areas, the alignment of speakers' vowel space areas was also taken into account. This is an important consideration to make, as two vowel spaces might have identical areas, but be different shapes, or show poor overlap. In either situation, the vowel spaces of the speakers could not be said to have aligned well, and consequently, it could be argued they still possess variation due to anatomical and physiological differences, and thus are unlikely to be directly comparable.

The alignment and overlap of vowel space areas under the different normalisation procedures were quantified and compared by the following method. Python v2.6.4 incorporating the Shapely v1.2.6 package was used to calculate and compute the areas of the intersection and union of all 20 speaker vowel space areas. Dividing the area of the intersection of all 20 speakers' vowel space areas by the area of the union of all 20 speakers' vowel space areas gave the percentage of area that overlapped. As the overall overlaps of the vowel spaces were calculated as percentages, they can be directly compared. A higher overlap percentage indicates better alignment of the vowel space areas by a normalisation procedure. Scale-drawn vowel space areas were also compared visually to confirm the results. Normalisation methods were again ranked according to how well they overall aligned the vowel space areas of the speakers.

6. Results

6.1. Equalising vowel space areas

Table 2 gives the SCV of speakers' hypothetical total vowel space areas under each normalisation method. The ranking of each method at effectiveness of equalising the vowel space areas based on these SCVs is also given.

As can be seen in Table 2, *Gerstman* displayed the smallest SCV of speaker vowel space areas, so can be said to have shown the least variance in vowel space area. It was thus most effective at equalising the vowel space areas. *LCE* was the next most effective, followed by *Lobanov* and *Bigham*.

1mW&F outperformed the original formulation, which in turn was more effective than *2mW&F* at equalising areas. *Nearey1* and *NeareyGM* performed the same with respect to equalising vowel space areas, giving no reason to favour one over the other based on this result alone. The exponential versions of Nearey's methods also gave identical results to one another, but were not as efficient at equalising the vowel space areas as the original formulations were.

Normalisation Method	SCV	Rank
Hertz	0.06212	N/A
Gerstman	0.01020	1
LCE	0.01487	2
Lobanov	0.02032	3
Bigham	0.02556	4
1mW&F	0.02587	5
Letter	0.02637	6
origW&F	0.02671	7
2mW&F	0.02818	8
ERB	0.03233	9
Nearey1	0.03250	=10
NeareyGM	0.03250	=10
Log	0.03250	=10
Ln	0.03250	=10
Bladon	0.03409	=14
Bark	0.03409	=14
Bark-diff	0.03549	16
Mel	0.03583	17
exp{Nearey1}	0.03798	=18
exp{NGM}	0.03798	=18
Nordström	0.03977	20

Table 2: SCVs of speakers' vowel space areas for each normalisation method.

All 20 normalisation techniques showed at least some improvement over raw Hertz values at equalising the vowel space areas of different speakers. *Nordström* performed the worst, followed by *exp{Nearey1}* and *exp{NGM}*, and then *Mel* and *Bark*. Although the three worst-performing methods were vowel-extrinsic, the results appear to imply that overall, vowel-intrinsic scaling formulae performed less well than vowel-extrinsic formulae.

6.2. Aligning vowel space areas

Table 3 presents the overlap percentages calculated and corresponding rankings awarded to each procedure based on their ability to align the speakers' vowel spaces. Impressionistically, the percentage overlaps fall into three distinct groups. Lines separating these groups have been drawn onto Table 3 to make the divisions more apparent.

Method	**Percent Overlapping**	**Rank**
Bigham	45.8%	1
2mW&F	43.8%	2
origW&F	43.4%	3
1mW&F	42.3%	4
Gerstman	30.0%	5
Lobanov	29.2%	6
Nordstrom	28.7%	7
exp{Nearey1}	27.6%	8
Nearey1	27.1%	9
exp{NGM}	26.9%	10
Bladon	25.9%	11
NeareyGM	25.7%	12
Letter	24.1%	13
LCE	23.1%	14
Bark-diff	13.5%	15
Bark	13.2%	16
Mel	13.1%	17
ERB	12.8%	18
Ln	12.2%	=19
Log	12.2%	=19
Hertz	12.6%	N/A

Table 3: Rankings for each normalisation method's ability to align speaker vowel spaces

Figure 3 displays the raw Hertz vowel space areas of the adult speakers. As can be seen, there is considerable difference in the sizes and shapes, with a clear distinction between the male and female speakers. As seen in Table 3, the percentage overlap was calculated as being just 12.6%.

Results for the vowel-intrinsic scaling methods were largely similar, with little if any overall improvement over raw Hertz in vowel space alignment, seen by overlap percentages increasing by less than 1%. As a visual illustration, Figure 4 gives the vowel space areas as normalised by *Bark*. The similarity to the raw Hertz areas is clearly evident.

The overlap percentages for *Log* and *Ln* were marginally smaller than for Hertz, suggesting that normalisation by these methods actually further misaligns speaker vowel spaces, albeit to a very minimal extent. *Bladon* was the best-performing vowel-intrinsic procedure, appearing mid-table. The total overlap of vowel spaces normalised by *Bladon* was nearly double that of those normalised using just *Bark*. *Letter* finished relatively low in the rankings, with an overlap percentage of 24.1%. This is quite surprising, as, visually, the vowel spaces appeared to align reasonably well, with one notable exception, as seen in Figure 5.

All four formulations of the Nearey method showed noticeable improvement over raw Hertz in aligning speaker vowel space areas, with *Nearey1* performing better than *NeareyGM*, and the exponential version of each method performing better than the original versions in each case, although there were only marginal differences in overlap percentages between the four renderings.

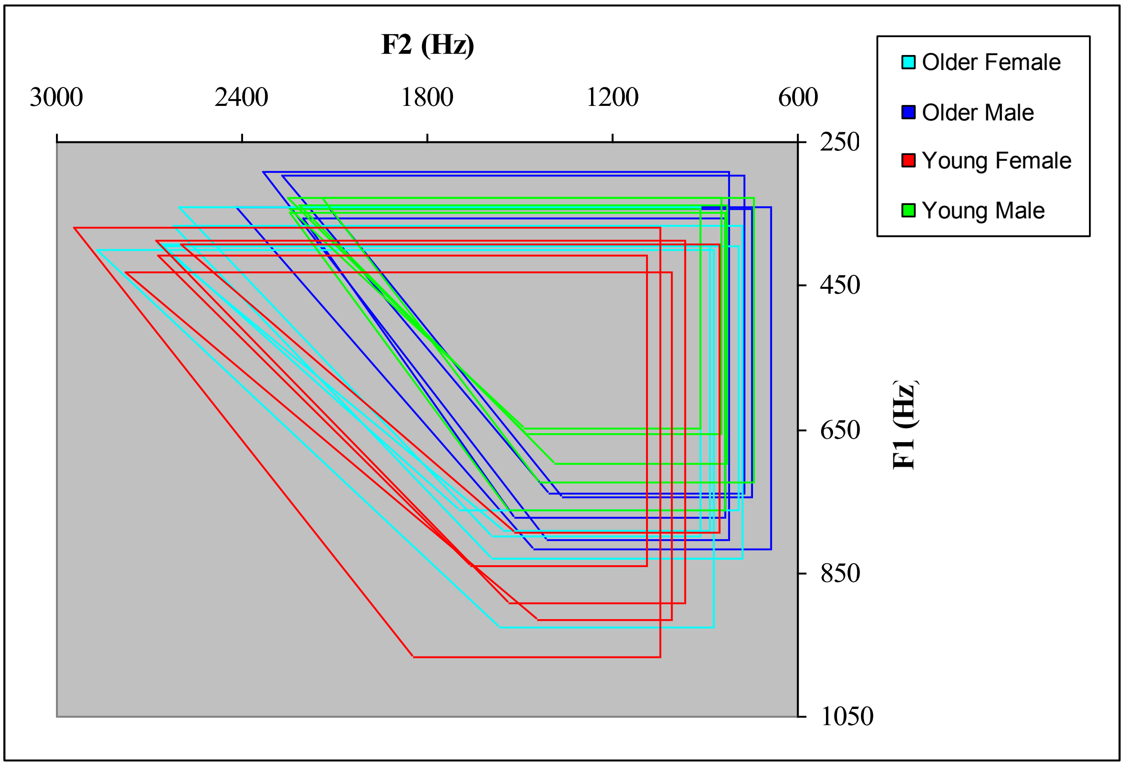


Figure 3: Raw Hertz vowel spaces of the 20 speakers

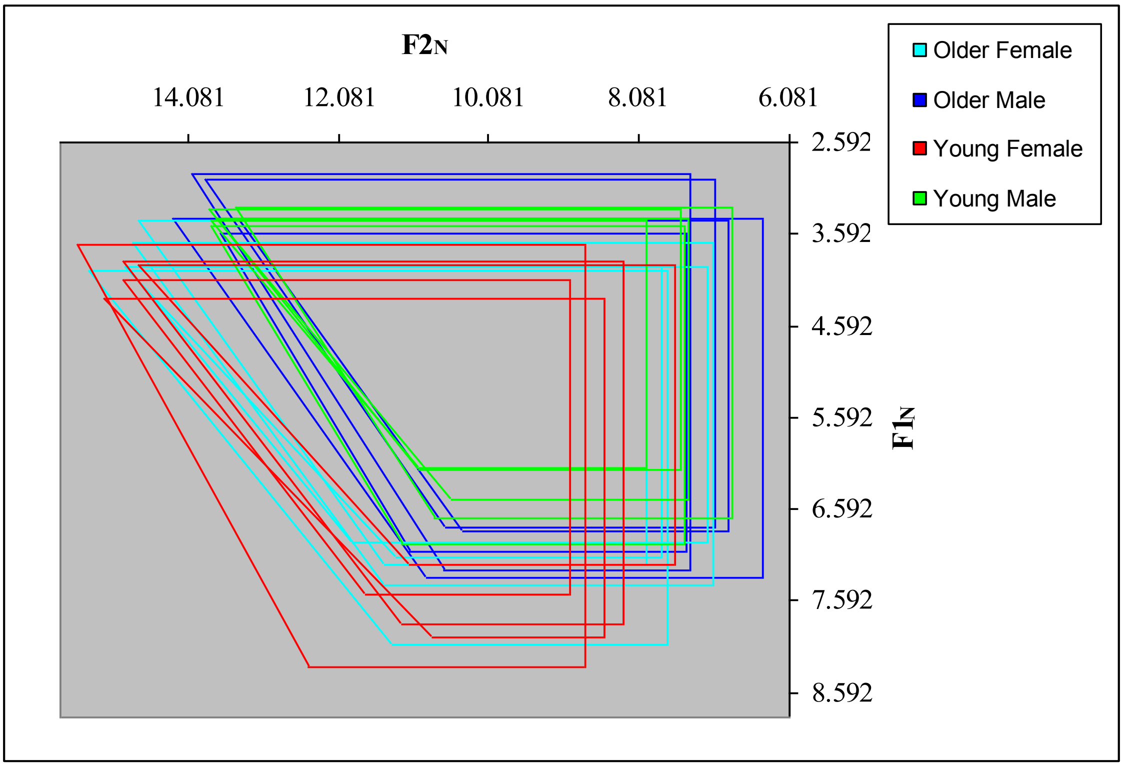


Figure 4: *Bark*-normalised vowel spaces of the 20 speakers

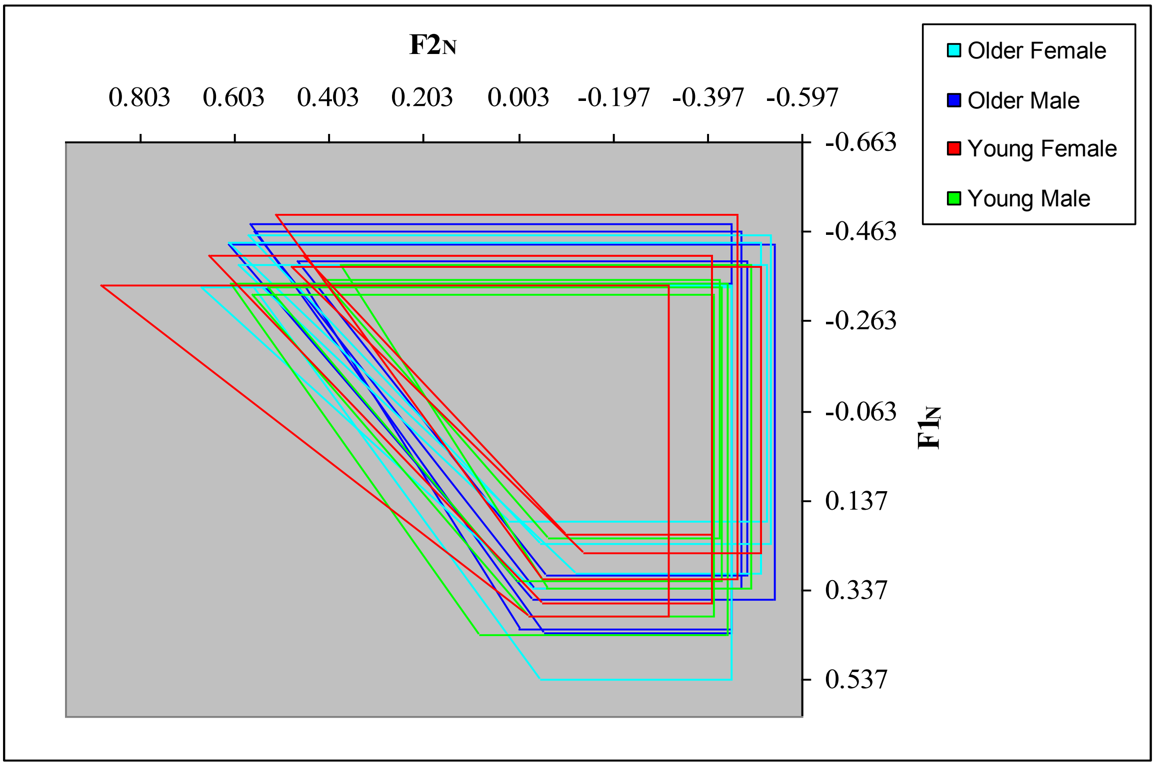


Figure 5: *Letter*-normalised vowel spaces of the 20 speakers

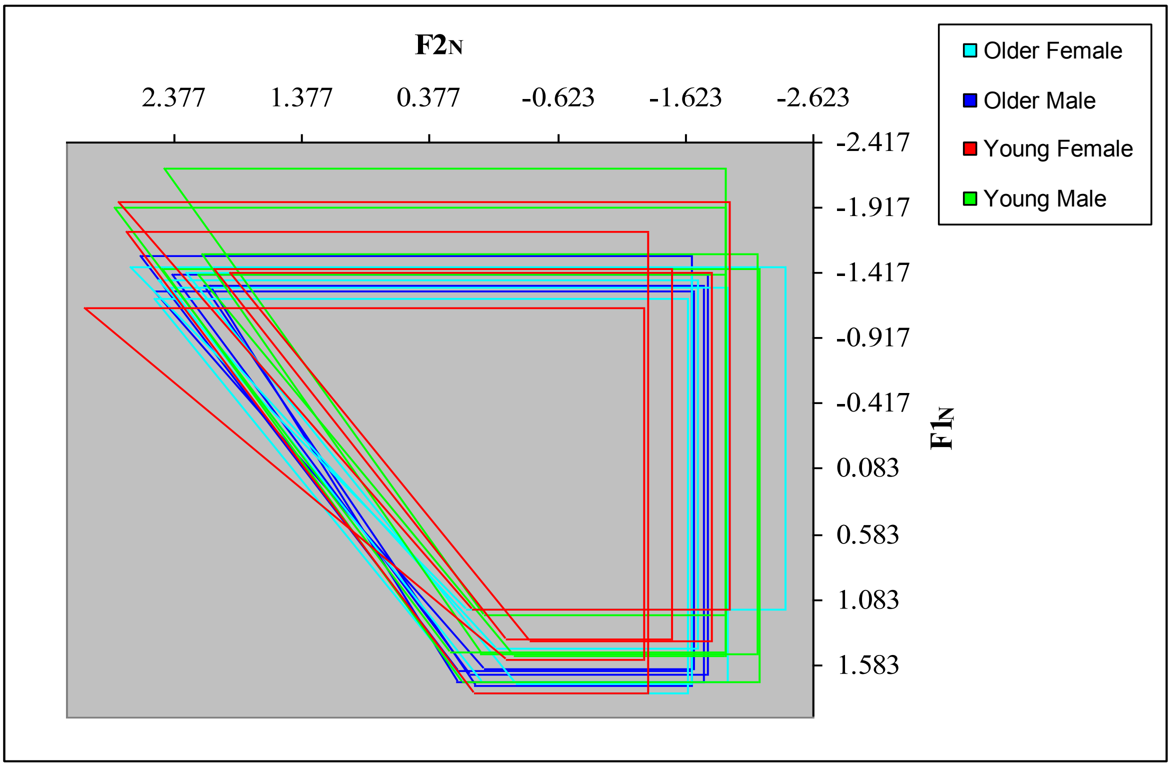
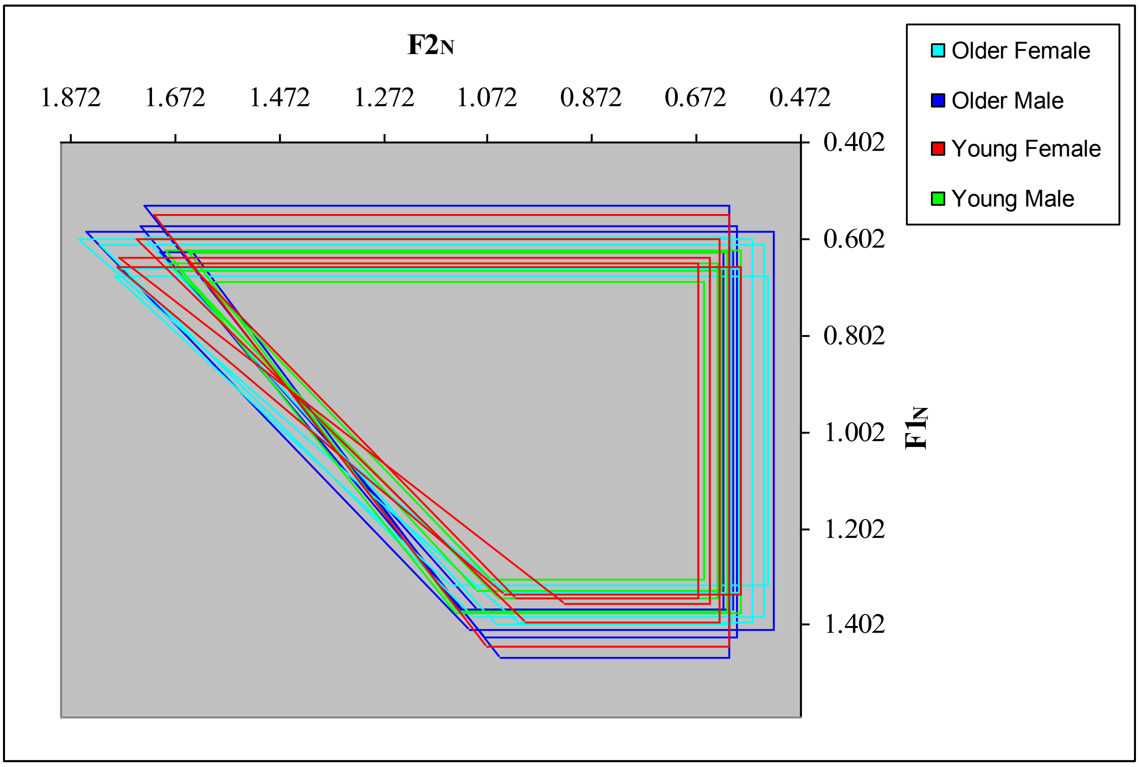


Figure 6: *Lobanov*-normalised vowel spaces of the 20 speakers

The overlap percentage for *Lobanov* was calculated to be 29.2%. This contrasts somewhat with the results of Fabricius et al. (2009: 427), who found the average overlap between a speaker's vowel space and the other speakers' vowel spaces collectively to be 56.4% for RP speakers, and as high as 68.8% for Aberdeen speakers for vowel spaces normalised using *Lobanov*. It is worth noting though, that the initial overlap percentages[5] for Fabricius et al.'s (2009) Hertz data are higher for both data sets than for this study, and the 16.6% overlap improvement by *Lobanov*-normalised data for this study is not that far removed from the 18.4% given by Fabricius et al. (2009: 427) for RP speakers. Inspection of the scale-drawn areas revealed that *Lobanov* failed to align vowel space areas of a number of the young speakers of both sexes, with mismatch at the top edge. However, looking beyond these speakers, the remaining vowel spaces appear to align fairly well as can be seen in Figure 6.

The three versions of *W&F* and *Bigham*, itself derived from *W&F*, were the most effective at aligning the vowel spaces. Indeed, there is a clear separation between these four procedures and the others with respect to overlap percentages, and all four showed vast improvement over raw Hertz measurements visually.

Bigham was ranked as being the procedure that performed the best at aligning the vowel space areas of the speakers, with an overlap percentage of 45.8%, more than triple that of the raw Hertz areas. Figure 7 clearly shows the high degree of alignment and overlap that normalising via *Bigham* facilitated, and comparison with Figure 3 (presented earlier) plainly shows this improvement visually. It is perhaps a little surprising that the overlap percentage isn't actually higher than 45.8% for *Bigham*. Certainly, the graphical results look to show greater overlap of the vowel space areas than this upon visual inspection. Nevertheless, it cannot be denied that increase to 45.8% from a baseline Hertz overlap percentage of 12.6% is a considerable improvement, albeit not to an idealised near-100% value.



[5] Fabricius et al. (2009) actually give their overlap calculations as decimals not percentages, but these have been converted to percentages here for direct comparison

Figure 7: *Bigham*-normalised vowel spaces of the 20 speakers

6.3. Overall Results

To complete the study, the results of the two tests were reviewed and aggregated to discover the overall relative performance of each procedure. A points system was developed with points awarded to procedures based on their rank position in the two comparisons. Points were then totalled and procedures were then ranked in order of points total with a lower score indicating a better result. Table 4 displays the points total and resulting overall ranking of each procedure, as well as summarising the previous results. The maximum possible points total was 40.

Normalisation Method	**Area SCV Reduction Rank**	**Area Aligning Rank**	**Total Points**	**Overall Rank**
Bigham	4	1	5	1
Gerstman	1	5	6	2
1mW&F	5	4	9	=3
Lobanov	3	6	9	=3
2mW&F	8	2	10	=5
origW&F	7	3	10	=5
LCE	2	14	16	7
Letter	6	13	19	=8
Nearey1	=10	9	19	=8
NeareyGM	=10	12	22	10
Bladon	=14	11	25	11
exp{Nearey1}	=18	8	26	12
ERB	9	18	27	=13
Nordström	20	7	27	=13
exp{NGM}	=18	10	28	15
Ln	=10	=19	29	=16
Log	=10	=19	29	=16
Bark	=14	16	30	18
Bark-diff	16	15	31	19
Mel	17	17	34	20

Table 4: Overall rankings for the 20 normalisation methods

7. Discussion

Consideration of the results brings to light some striking patterns and correlations, some in accordance with what other comparative studies of normalisation procedures have found, and some in apparent contrast.

The five vowel-intrinsic scaling transformations all performed poorly in the comparisons conducted, as did *Bark-diff*, itself a vowel-intrinsic method based on manipulations of *Bark*-transformed data. *Mel* was the worst-performing normalisation technique overall, followed by

Bark-diff and *Bark*. This result, coupled with similar findings by Adank (2003), Adank et al. (2004) and Clopper (2009) confirms that vowel-intrinsic methods are less than adequate for the purposes of vowel formant normalisation, for a sociophonetic study at least. The point should be made, however, that even the worst-performing procedure, *Mel*, was an improvement on the raw Hertz frequency measures, suggesting that any form of normalisation is better than not normalising at all. It is possible that vowel-intrinsic normalisation has a role to play in studies of human vowel cognitive perception, but with respect to sociophonetic research, their relative ineffectiveness to more superior methods makes their use redundant.

All formulations of Nearey's method performed relatively poorly in the comparative tests, and, thus, overall. This adds to recent literature that has reported outperformance of Nearey's method by other procedures. For example, Langstrof (2006) ranked *NeareyGM* less favourably than other procedures he tested, while Fabricius et al. (2009) showed *Nearey1* was statistically significantly outperformed by *Lobanov*, *1mW&F* and *origW&F*. In contrast, other researchers have ranked one or other of Nearey's methods as performing well (Clopper 2009; Adank et al. 2004; Adank 2003; Disner 1980; Hindle 1978). Indeed, Disner (1980) concluded that *Nearey1* was the most successful of the methods she tested at normalising effectively, while the same procedure was found by Adank et al. (2004) to be joint best at removing gender-related variation assumed to correspond to anatomical differences. Both of these results are in contrast to the findings of this study.

When considering the four different versions of Nearey's method that were compared, the results of these comparative tests showed that treatment of the formant values was not identical. The conclusion can be drawn, that use of an exponential Nearey formula rather than an original formulation does affect the outcome of the normalised data. Moreover, based on the results of these two tests, *exp{Nearey1}* and *exp{NGM}* both performed worse than their original counterpart. Use of a grand log-mean versus individual log-means was also observed to have an effect on the results of the normalisation. *Nearey1* was found to perform marginally better than *NeareyGM*. This finding corresponds with results of Adank (2003) and Adank et al. (2004).

Letter and *LCE* both performed well at equalising speaker vowel space areas, but comparatively poorly at aligning the vowel space areas, while *Nordström* and *exp{Nearey1}* were both far better vowel space aligners than equalisers. These results demonstrate the possibility of procedures performing to different levels of effectiveness depending on the method of comparison used, and suggest evaluation of procedures should ideally be based on a range of comparative tests.

There is a clear separation between the procedures ranked in the top six overall and the others. Of the six, *origW&F* and *2mW&F* performed least well overall. In correspondence with findings by Fabricius et al. (2009), *1mW&F*, which (following comments by Thomas & Kendall 2007) does not use the F_2 value of the point [a] when constructing the centroid *S* used to normalise data, outperformed *origW&F*. Furthermore, it performed strongly in both comparisons, and so ought definitely to be at least considered when choosing a normalisation procedure. In keeping with Fabricius et al.'s (2009) findings, *1mW&F* performed equally as well as *Lobanov*.

Lobanov, finishing in equal third place, is perhaps not ranked as highly here as in other comparative studies. The inclusion of additional methods not compared by other researchers may be related to this. For example, Adank (2003), who found *Lobanov* to be the best procedure overall, did not include in her study any formulation or derivation of the Watt & Fabricius method, two of which (*1mW&F* and *Bigham*) finished as high as, or higher than, *Lobanov* in the overall rankings of this experiment. It should be noted that *Lobanov* still performed well, and considerably better than the majority of the procedures tested. Its existing widespread use in the sociolinguistic world is, therefore, warranted.

A method that performed strongly in both comparative tests but has been largely ignored as a viable procedure by studies thus far, was *Gerstman*. Langstrof (2006), Adank et al. (2004), Adank (2003) and Clopper (2009) all ranked *Gerstman* highly in terms of performance, and remarked on its effectiveness. This combined evidence points to *Gerstman* being an arguably adequate choice for a formant normalisation technique.

Based on the results of equalising and aligning the vowel space areas of different speakers, *Bigham* was evaluated as performing marginally the best overall, closely followed by *Gerstman*. The downside to both these procedures is that neither is available as part of NORM, therefore they are not as easily accessible or useable by researchers as other methods. Of the procedures available in NORM, *1mW&F* and *Lobanov* were the best-performing.

8. Conclusion

This chapter has provided an overview of the available and existing vowel formant normalisation procedures, and compared their effectiveness at normalising formant frequencies from multiple speakers of both sexes and of two age groups.

The research conducted extends the existing normalisation literature as it considers a large dataset consisting of large numbers of tokens from a wide range of vowels from a relatively large speaker sample. The data used were collected for a sociophonetic study, and therefore are neither laboratory-controlled, nor artificially created, and were recorded under conditions typical of sociolinguistic research, meaning the results are not solely applicable to speech data recorded under laboratory conditions.

20 different normalisation procedures were compared to give as full a picture as possible about their relative effectiveness. It is believed that the only established procedures not included are those of Miller (1989) and Labov et al. (2006). Miller's method was omitted as fundamental frequency measurements are needed to perform the formula, and these had not been made.

The decision was made to exclude Labov et al.'s method because the number of speakers used, 24, was deemed insufficient to give robust results. Thomas & Kendall (2007) note that Labov et al.'s procedure is a viable possibility only when the number of speakers is 'exceptionally high'. A figure of 345 speakers is posited (Thomas & Kendall 2007), far exceeding the 20 used for this study.

It is acknowledged that some of the methodology used as part of this study could be criticised. Jacewicz et al. (2007) found that a pentagon better estimated the complete vowel space area used by speakers, while Fox & Jacewicz (2008) hypothesised that a polygon defined by joining all perimeter vowels including diphthongs should be used to avoid underestimating the working space of the vowel system. Bearing these points in mind, the decision to construct and define vowel space areas as quadrilateral for the comparative tests involving equalising and aligning vowel spaces could be questioned. However, it should be noted that the method was used consistently, so the areas defined, though possibly not the exact total vowel space for a speaker, were still directly comparable as they were constructed in the same way for all speakers. Principally, the end results of the comparative tests produced robust results analogous to those of similar existing studies.

The nine best-performing methods all showed the typological classification of vowel-extrinsic, formant-intrinsic, speaker-intrinsic. It could be argued that this is abundant evidence in support of the hypothesis made by Adank (2003) and Adank et al. (2004), and recounted by Clopper (2009) and Fabricius et al. (2009), among others, that it is this type of normalisation method that performs best and is the most suitable to use for language variation research. In comparison, the poor performance by vowel-intrinsic methods, replicating findings by Adank (2003), Adank et al. (2004) and Clopper (2009), imply they are less than adequate for use in sociophonetic research.

Acknowledgements

Thank you to Paul Foulkes, Dom Watt, Bill Haddican, Dan Ezra Johnson, Greg Flynn and two anonymous reviewers for helpful comments and feedback. Thanks are also due to Jess Wardman and James Porter for help with Shapely and Python coding, and Huw Llewelyn-Jones for computer-related assistance. The Praat script used for formant extraction was devised by Phil Harrison of J.P.French Associates. Research was funded by the ESRC.

References

ADANK, PATTI. 2003. *Vowel Normalisation: A Perceptual-Acoustic Study of Dutch Vowels*. PhD Dissertation. Nijmegen: University of Nijmegen.

ADANK, PATTI, SMITS, ROEL & VAN HOUT, ROELAND. 2004. "A comparison of vowel normalisation procedures for language variation research". *Journal of the Acoustical Society of America* 116(5), 3099-3107.

BEAL, JOAN C. 2008. "English dialects in the North of England: Phonology". In: Bernd Kortmann & Clive Upton (eds) *Varieties of English. Vol. 1: The British Isles*. Berlin: Mouton de Gruyter. pp. 122-144.

BIGHAM, DOUGLAS S. 2008. *Dialect Contact and Accommodation among Emerging Adults in a University Setting*. PhD Dissertation. Austin: University of Texas at Austin.

BLADON, R.ANTHONY W., HENTON, CAROLINE G. & PICKERING, J.BRIAN. 1984. "Towards an auditory theory of speaker normalisation". *Language and Communication* 4(1), 59-69.

CLARK, URSZULA. 2008. "The English West Midlands: Phonology". In: Bernd Kortmann & Clive Upton (eds) *Varieties of English. Vol. 1: The British Isles*. Berlin: Mouton de Gruyter. pp. 145-177.

CLOPPER, CYNTHIA G. 2009. "Computational methods for normalising acoustic vowel data for talker differences". *Language and Linguistic Compass* 3(6), 1430-1442.

DISNER, SANDRA F. 1980. "Evaluation of vowel normalisation procedures". *Journal of the Acoustical Society of America* 67(1), 253-261.

FABRICIUS, ANNE H. 2008. *Vowel Normalisation in Sociophonetics: When, Why, How?* Paper presented at Sociolinguistics Circle, Copenhagen University, 16th September 2008.

FABRICIUS, ANNE H., WATT, DOMINIC J.L. & JOHNSON, DANIEL E. 2009. "A comparison of three speaker-intrinsic vowel formant frequency normalisation algorithms for sociophonetics". *Language Variation and Change* 21(3), 413-435.

FLYNN, NICHOLAS E.J. 2007. *A Sociophonetic Comparison of Adolescent Speakers in two Areas of Nottingham.* MA Dissertation. Colchester: University of Essex.

FLYNN, NICHOLAS E.J. fc. *Levelling and Diffusion at the North/South Border: A Sociophonetic Study of Nottingham Speakers* [working title]. PhD Dissertation. York: University of York.

FOX, ROBERT A. & JACEWICZ, EWA. 2008. “Analysis of total vowel space areas in three regional dialects of American English”. In: *Proceedings of Acoustics ’08 Paris*. pp. 495-500.

GERSTMAN, LOUIS. 1968. “Classification of self-normalised vowels”. *IEEE Transactions of Audio Electroacoustics* AU-16. pp.78-80.

GLASBERG, BRIAN R. & MOORE, BRIAN C.J. 1990. “Derivation of auditory filter shapes from notched noise data”. *Hearing Research* 47(1-2), 103-138.

HINDLE, DONALD. 1978. ‘Approaches to vowel normalisation in the study of natural speech”. In: David Sankoff (ed) *Linguistic Variation: Models and Methods*. New York: Academic Press. pp. 161-171.

INTERNATIONAL PHONETIC ASSOCIATION. 1999. *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

JACEWICZ, EWA, FOX, ROBERT A. & SALMONS, JOSEPH. 2007. “Vowel space areas across dialects and gender”. In: J. Trouvain & W. Barry (eds) *Proceedings of the 16th International Congress of Phonetic Sciences*, 6th – 10th August 2007, Saarbrücken. Saarbrücken: Universität des Saarlandes. pp. 1465-1468.

JOHNSON, KEITH. 2003. *Acoustic and Auditory Phonetics*. (2nd edn) Oxford: Blackwell.

KAMATA, MIHO. 2008. *An Acoustic Sociophonetic Study of three London Vowels*. PhD Dissertation. Leeds: University of Leeds.

LABOV, WILLIAM, ASH, SHARON & BOBERG, CHARLES. 2006. *The Atlas of North American English: Phonetics, Phonology and Sound Change*. Berlin: Mouton de Gruyter.

LANGSTROF, CHRISTIAN. 2006. *Vowel Change in New Zealand English – Patterns and Implications*. PhD Dissertation. Christchurch, New Zealand: University of Canterbury.

LOBANOV, B.M. 1971. “Classification of Russian vowels spoken by different speakers”. *Journal of the Acoustical Society of America* 49(2), 606-608.

MILLER, JAMES D. 1989. “Auditory-perceptual interpretation of the vowel”. *Journal of the Acoustical Society of America* 85(5), 2114-2134.

NEAREY, TERRANCE M. 1978. *Phonetic Feature Systems for Vowels.* PhD Dissertation reproduced by the Indiana University Linguistics Club. Indiana: Indiana University Linguistics Club.

NORDSTRÖM, PER-ERIK. 1977. “Female and infant vocal tracts simulated from male area functions”. *Journal of Phonetics* 5(1), 81-92.

ROSNER, BURTON S. & PICKERING, J.BRIAN. 1994. *Vowel Production and Perception*. Oxford: Oxford University Press.

STEVENS, STANLEY S. & VOLKMANN, JOHN. 1940. “The relation of pitch to frequency: A revised scale”. *American Journal of Psychology* 53(3), 329-353.

SYRDAL, ANN K. & GOPAL, H.S. 1986. “A perceptual model of vowel recognition based on the auditory representation of American English vowels”. *Journal of the Acoustical Society of America* 79(4), 1086-1100.

THOMAS, ERIK R. 2002. “Instrumental phonetics”. In: J.K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds) *The Handbook of Language Variation and Change*. Oxford: Blackwell. pp. 168-200.

THOMAS, ERIK R. & KENDALL, TYLER. 2007. *NORM: The Vowel Normalisation and Plotting Suite*. Online Resource. URL: <http://ncslaap.lib.ncsu.edu/tools/norm> Accessed: 17/11/2008.

TOLLFREE, LAURA F. 1999. “South East London English: Discrete versus continuous modelling of consonantal reduction”. In: Paul Foulkes & Gerard J. Docherty (eds) *Urban Voices: Accent Studies in the British Isles*. London: Arnold. pp. 163-184.

TRAUNMÜLLER, HARTMUT. 1990. "Analytical expressions for the tonotopic sensory scale". *Journal of the Acoustical Society of America* 88(1), 97-100.
TRUDGILL, PETER. 1986. *Dialects in Contact*. Oxford: Blackwell.
WATT, DOMINIC J.L. & ALLEN, WILLIAM H.A. 2003. "Illustrations of the IPA: Tyneside English'. *Journal of the International Phonetic Association* 33(2), 267-271.
WATT, DOMINIC J.L. & FABRICIUS, ANNE H. 2002. "Evaluation of a technique for improving the mapping of multiple speakers' vowel spaces in the F_1~F_2 plane". *Leeds Working Papers in Linguistics and Phonetics* 9, 159-173.
WATT, DOMINIC J.L., FABRICIUS, ANNE H. & KENDALL, TYLER. 2010. "More on vowels: Plotting and normalising". In: Marianna di Paolo & Malcah Yaeger-Dror (eds) *Sociophonetics: A Student's Guide*. London: Routledge. pp. 107-118.
WATT, DOMINIC J.L. & MILROY, LESLEY. 1999. "Patterns of variation and change in three Newcastle vowels: Is this dialect levelling?". In: Paul Foulkes & Gerard J. Docherty (eds) *Urban Voices: Accent Studies in the British Isles*. London: Arnold. pp. 25-46.
WELLS, J.C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Appendix

Normalisation Method	**Equation(s) Used**
Log	$F_i^N = \log_{10}(F_i)$
Ln	$F_i^N = \ln.(F_i)$
ERB	$F_i^N = 21.4 \ln.(0.00437F_i + 1)$
Mel	$F_i^N = 1127 \ln.\left(1 + \frac{F_i}{700}\right)$
Bark	$F_i^N = 26.81 \left(\frac{F_i}{1960 + F_i}\right) - 0.53$
Bladon	$F_i^N = \begin{cases} 26.81\left(\frac{F_i}{1960 + F_i}\right) - 0.53 & \text{, speaker is male} \\ \left(26.81\left(\frac{F_i}{1960 + F_i}\right) - 0.53\right) - 1 & \text{, speaker is female} \end{cases}$
Bark-diff	$F_i^N = B_3 - B_i$, $i < 3$ (B_i = Bark-transformed F_i)
LCE	$F_i^N = \frac{F_i}{F_i^{\max}}$

Gerstman	$F_i^N = 999\left(\dfrac{F_i - F_i^{min}}{F_i^{max} - F_i^{min}}\right)$
Lobanov	$F_i^N = \dfrac{(F_i - \mu_i)}{\sigma_i}$
Nordström	$F_i^N = \begin{cases} F_i & \text{, speaker is male} \\ \left(\dfrac{\mu_{F_3}^{male}}{\mu_{F_3}^{female}}\right) F_i & \text{, speaker is female} \end{cases}$ where $\dfrac{\mu_{F_3}^{male}}{\mu_{F_3}^{female}} = 0.876\ldots$ for this dataset.
origW&F	$F_i^N = \dfrac{F_i}{S(F_i)}$ $S(F_i) = \dfrac{F_i[i] + F_i[a] + F_i[u']}{3}$ $F_i[i] = F_i[\text{FLEECE}]$ $F_i[a] = F_i[\text{TRAP}]$ $F_1[u'] = F_2[u'] = F_1[i]$
1mW&F	$F_i^N = \dfrac{F_i}{S(F_i)}$ $S(F_i) = \begin{cases} \dfrac{F_i[i] + F_i[a] + F_i[u']}{3} & , i = 1 \\ \dfrac{F_i[i] + F_i[u']}{2} & , i = 2 \end{cases}$ $F_1[i] = F_1^{min}$, $F_2[i] = F_2^{max}$, $F_1[a] = F_1^{max}$, $F_1[u'] = F_2[u'] = F_1[i]$
2mW&F	$F_i^N = \dfrac{F_i}{S(F_i)}$ $S(F_i) = \begin{cases} \dfrac{F_i[i] + F_i[a] + F_i[u']}{3} & , i = 1 \\ \dfrac{F_i[i] + F_i[u']}{2} & , i = 2 \end{cases}$ $F_1[i] = F_1^{min}$, $F_2[i] = F_2^{max}$, $F_1[a] = F_1^{max}$, $F_1[u'] = F_1^{min}$, $F_2[u'] = F_2^{min}$
Bigham	$F_i^N = \dfrac{F_i}{S(F_i)}$ $S(F_i) = \dfrac{F_i[i'] + F_i[a'] + F_i[o'] + F_i[u']}{4}$ $F_1[i'] = F_1^{min}$, $F_2[i'] = F_2^{max}$, $F_1[a'] = F_1^{max}$, $F_2[a'] = F_2[\text{TRAP}]$, $F_1[o'] = F_1^{max}$, $F_2[o'] = F_2^{min}$, $F_1[u'] = F_1^{min}$, $F_2[u'] = F_2^{min}$

Letter	$F_i^N = \dfrac{F_i}{F_i[\mathit{lett}\text{ER}]} - 1$

Nearey1	$F_i^N = \ln.(F_i) - \mu_{\ln.(F_i)}$
NeareyGM	$F_i^N = \ln.(F_i) - \mu_{\ln.(F_j)} \quad , \ \forall\, j = 1, \ldots, n$
exp{Nearey1)	$F_i^N = \exp.\{\ \ln.(F_i) - \mu_{\ln.(F_i)}\}$
exp{NGM}	$F_i^N = \exp.\left\{\ln.(F_i) - \left(\dfrac{\mu_{\ln.(F_1)} + \mu_{\ln.(F_2)}}{2}\right)\right\} \quad , \ \forall\, j = 1, \ldots, n$

Nicholas Flynn
Department of Language and Linguistic Science
University of York
Heslington
York
YO10 5DD
email: nejf100@york.ac.uk

OLD ENGLISH PSYCH VERBS AND QUIRKY EXPERIENCERS

LUIZ GUILHERME GUIDI

Abstract

This paper presents a corpus–based study of Old English psychological verbs (e.g. *andian* ‘to envy’, *abelgan* ‘to annoy’), examining them based on recent theoretical advances on the syntax of this class of verbs (Landau, 2009) and presenting the occurrence of reflexive psych verbs (e.g. *(a(n)/on–)drædan* ‘to fear’), a pattern not previously noted in the literature. From this new perspective, the data suggests that there is some evidence supporting the claim that Old English non–nominative Experiencers (accusative and dative) can be clausal subjects i.e. quirky subjects (in parallel with Allen, 1995). The finding impacts on the debate regarding whether topicalisation in Old English was restricted to main clauses or also allowed in subordinate clauses (van Kemenade, 1987, 1997; Pintzuk, 1999). The paper also discusses the difficulties of establishing the unacusativity of certain Old English psych verbs.

1. Introduction

This papers aims to: i) give a description of Old English (OE) psych verbs and provide a syntactic account on the basis of a new theory of the syntax of psych verbs (Landau, 2009); ii) examine data to test predictions made from (i) that OE non–nominative Experiencers may be quirky subjects. In addition, the corpus–data surveyed reveals the existence of reflexive psych verbs in OE which had not been previously noted and is exemplified below.

(1)	*and*	*feawa*	*him*	*ondrædað*	*þære*	*sawle*	*deað.*
	and	few.nom	themselves.dat	fear	the	soul.gen	death.acc

‘and few people fear the death of the soul’

(coaelhom,+AHom6:143.948)

It has been often assumed that OE psych verbs are unaccusative and that has been used to explain the common occurrence of preposed non–nominative Experiencers (van Kemenade, 1997). It will be shown that determining the unaccusativity of non–nominative Experiencer verbs cannot be firmly determined and, more importantly, that OE non–nominative Experiencers may be (quirky) subjects, allowing them to be fronted. The example below demonstrates this, with the dative Experiencer *þam gifran* as the clausal subject.

(2)	*gif*	*þam*	*gifran*	*angemetlic*	*spræc*	*ne*	*eglde*
	If	the	greedy.dat	eloquent	speech.nom	not	afflicted

‘if the greedy are not afflicted by loquacity’

(cocura,CP:43.309.2.2069)

From a typological perspective, this would suggest that OE patterns with Scandinavian languages such as Icelandic with respect to what kinds of quirky subjects are allowed.

A brief discussion of psych verbs is given in 2, together with the main theoretical assumptions of the paper. Following that, section 3 introduces the theory of psych verbs proposed by Landau (2009) that will serve as basis to analyse the OE data presented in section 4. The analysis is presented in section 5 and in section 6, I will present and discuss evidence that OE dative Experiencers may be quirky subjects. Section 7 is the conclusion.

2. Background

Psych verbs are verbs that express (a change in) mental states and a relation between an Experiencer and the subject or cause of such psychological state and have been the focus of much debate in the theoretical syntax literature (Belletti and Rizzi, 1988; Pesetsky, 1995; Anagnostopoulou, 1999; Landau, 2009). A number of studies have revealed that, while verbs with an Experiencer subject have largely uncontroversial syntax, those with an Experiencer in object position behave unpredictably, exhibiting a number of seemingly conflicting phenomena (see Landau, 2009). The sentences below are typical examples of psych verbs.

(3) The Afghan people fear the British army

(4) The terrorist attacks frightened the civilians

With specific reference to OE, the debate on psych verbs has focused mainly on diachronic change and their status as impersonal verbs (Fischer and van der Leek, 1983; Denison, 1993; Allen, 1995; Fischer et al., 2000, amongst others). The term 'impersonal' has often been used to refer to psych verbs in the OE literature at least since Jespersen (1984) and van der Gaaf (1904), even though not all psych verbs give rise to subjectless constructions (typically, a clause that lacks a nominative subject).

(5)	*þæt*	*him*	*on*	*sumne*	*sæl*	*huru*	*gesceamige*	*hyra*	*stuntnyssa*
	that	him.dat	on	some	occasions	even	embarasses	his	foolishness.gen

'that on some occasions he is ashamed of his foolishness'

(colwstan1,+ALet2[Wulfstan1]:6.9)

(6)	*þæt*	*he*	*men*	*geswence*	*for*	*heora*	*misdædum*
	that	he.nom	men.dat	irritated	for	their	misdeed

"that he irritated men for their misdeed'

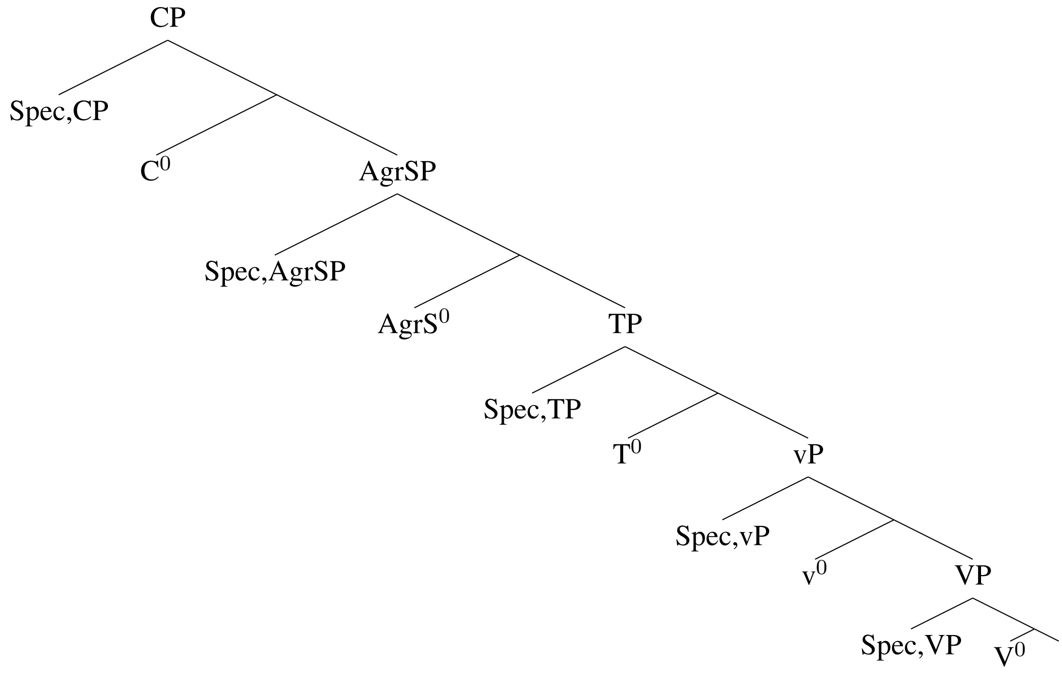
(coaelive,+ALS[Auguries]:177.3607)

Both examples above have been associated with impersonal constructions but while sentence (5) is a true impersonal construction, with no nominative argument, (6) has a nominative subject *he* and is not impersonal. To avoid confusion, I will use the term psych verbs to refer to these verbs rather than impersonal.

Many analyses have been put forward to account for these facts but I believe they lack a more solid and precise characterisation of their syntax as only recently syntacticians started to understand their properties. Therefore, this paper focuses on giving a synchronic account of OE transitive psych verbs with the framework of Landau (2009) revealing patterns not previously discussed in the literature.

This paper adopts the syntactic framework of the Minimalist Program (Chomsky, 1995) largely based on Adger (2003), where syntactic derivations are the result of the recursive operation Merge which joins two syntactic elements as specified by their selectional requirements. The approach taken here differs from that of Adger to follow more recent formulations where head-movement takes place at PF and phrase movement is driven by the requirement of checking the EPP feature of a head like C or or *v* (Chomsky, 2001). In addition, following Haeberli (2000) I assume that the OE IP (or TP for Adger) is split between the heads AgrS and T, resulting in the core structure below.[1]

(7)



With respect to Θ–roles, I adopt a weaker version of argument linking principles where arguments are mapped onto the structure according to an ordering hierarchy (Grimshaw, 1990) rather than being associated with specific structural positions (Adger follows Baker's (1988) UTAH). This difference is required to make the syntax compatible with the theory of psych verbs to be presented in section 3. In addition, arguments can check their case features either in the syntactic structure by a functional head (I checks nominative [nom] and v accusative [acc]) or in parallel with Θ–role assignment, in which case an argument's case marking is inherent or lexical (as in Chomsky (2000)).[2]

Here I follow Haeberli (2000) and assume that the verb second phenomenon (V2) in OE main clauses is a result of V–movement to C or to AgrS, depending on the clause–type, and that Spec,TP is the position for non–proniminal subjects (based on van Kemenade (1987, 1997, hereafter VK) Pintzuk (1993, 1995, 1999, hereafter SP)),[3] with the EPP feature of T imposing the requirement for subject position always to be filled. Following SP, I also assume that OE VPs may vary between VO (verb–object) and OV (object–verb) ordering due to synchronic phrase structure competition, and that the same variation is true for Infl (AgrS and T).

3. *The Locative Syntax Of Experiencers*

3.1. *Experiencers as mental locations*

This section introduces Landau's (2009) recent theory of psych verbs *The Locative Syntax of Experiencers* (henceforth LSE), where he draws on evidence from a variety of languages and sources accumulated over the years and presents an elegant account that will serve as the basis for my analysis of OE psych verbs.

The keystone of the proposal is that Experiencers are (mental) locations. He argues that, on the assumption that language and thought are cognitively related, Experiencers are conceptually encoded as "mental locations – containers or destinations of mental states/effects" (Landau, 2005, 7). What this means is that Experiencers are essentially locatives, in the sense of receivers of experience. and a sentence like (8a) can have the conceptual representation schematically defined as (8b).

(8) **a** X frightens Y

b $[CS^{+}([X]^{\alpha}, [INCH [BE ([FEAR ([\alpha]), [AT [Y]]])]])$[4]

c X causes fear of X to come to be in Y

(Jackendoff (1990, 300) in Landau (2005, 7))

In some cases, the locative nature of Experiencers is linguistically transparent as in the English and Hebrew examples below.

(9) There is in me a great admiration for painters

(Arad (1998, 228) in Landau (2005, 9))

(10) *yeš be-Gil eyva gdola klapey soxney bituax.*
there-is in-Gil rancor great toward agents-of insurance
'Gil has a great rancour toward insurance agents'

(Landau, 2005, 9)

As such, Experiencers should share many of the properties of true locatives and Landau argues that this is manifested through the following syntactic properties:

(11) All object Experiencers are oblique (or dative)

(12) Experiencers are LF–subjects

(Landau 2005:5)

Given the locative nature of Experiencers, the first property is justified by the fact that a preposition normally introduces non-subject locatives and, as a consequence of it, Landau assumes that all object Experiencers are complements of a preposition, overt or null, and that non–nominative Experiencers always bear inherent case.

Furthermore, in parallel with the subjecthood of locatives as in locative inversion (see Bresnan, 1994), all Experiencers are subjects, be it overtly or at LF. The proposal is that the head T encodes spatio–temporal features and the locative (i.e. spatial) character of Experiencers requires them to be locally related to T, triggering their movement to Spec,TP. While this is true for subject Experiencers by default (they are always overt nominative subjects), in LSE 'object' Experiencers are also subjects, varying cross–linguistically with respect to whether the subjecthood requirement is satisfied before Spell–Out or at LF. This restriction is dependent on the case of the Experiencer and gives rise to three main typological patterns:

(13) Possible Case of Quirky Subjects

a All cases: Icelandic, Faroese, Greek.

b Dative only: Italian, Spanish, Dutch.

c No case: English, French, Hebrew.

(Landau, 2005, 80)

In order to account for the differences between psych verbs, Landau divides them into three types, largely following Belletti and Rizzi's (1988) classic paper. The classification is presented next, along with their syntactic properties.

3.2. Subject Experiencer verbs: Class I

As mentioned above, the syntax of these verbs is fairly transparent. Verbs falling under this class are those such *love*, *fear* and *hate* which select an Experiencer and a Theme case marked as below:[5]

(14) Bill loves Mary (Experiencer[nom], Theme[acc])

Verbs of this class have the Experiencer argument projected as an external argument which is, consequently, the nominative subject, and [acc] is checked on Theme by v.[6] The LF–subjecthood requirement of Experiencers is always satisfied before Spell–Out and, thus, overt.

3.3. Object Experiencer verbs: Class II

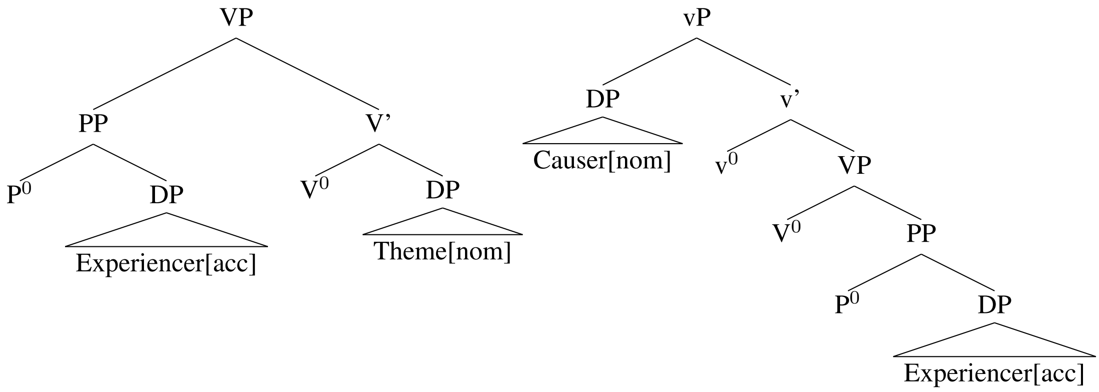
Class II is the most controversial of the three verb classes. According to LSE, the difficulty comes from the fact that most of the verbs that fall under this category are ambiguous between an eventive and a stative reading.

(15) Mr Brown annoyed the voters unintentionally\on purpose (stative\eventive)

The aspectual distinction between the sentences above is a result of their different grammatical structure: in both uses *the voters* is the Experiencer, but the subject *Mr Brown* is a Theme in the stative reading (with *unintentionally*) and a Causer in the eventive one (*on purpose*).

The dissociation in θ–role assignment also reflects their unaccusativity. Following well–established principles of lexical projection, Causers are merged by v, thus, externally; on the other hand, Themes are projected VP–internally. Following (11), Experiencers always bear inherent accusative case and are introduced by a preposition. Given that and the assumption in LSE that arguments bearing inherent case are internal arguments, only Class II–statives are unaccusatives, for they have no external argument. Nominative case in the Theme or Causer argument is checked by [nom] in T. For clarity, their basic structure is demonstrated below.

(16) a. Stative b. Eventive



With regards to the subjecthood requirement, their accusative case means that a quirky Experiencer subject with these verbs is only possible in languages that fall under the typology (13a), as in the Icelandic example below.

(17) *Mig dreymdi mmu.*
I.acc dreamt grandma.acc
'I dreamt of grandma'

(Barðal (1999, 4) in Landau (2005, 77))

3.4. Object Experiencer verbs: Class III

Class III verbs have a very similar syntax to Class II–statives differing only in the inherent case of the Experiencer, which is dative. In ModE, the preposition that selects it is overt.

(18) The idea appealed to Bill (Theme[nom], Experiencer[dat])

According to Landau, preposed dative Experiencers are well attested phenomenon in typologically unrelated languages such as Greek, Dutch, Italian and many others. He notes that in many of these languages it is possible to alternate the ordering of the arguments (the same is true for Class II verbs) and that, according to Mulder (1992), dative first order is unmarked in Dutch.

(19) *dat de taalkundige die analyse opviel*
that the linguist.dat the analysis.nom occurred-to

(20) *dat die analyse de taalkundige opviel*
that the analysis.dat the linguist.nom occurred-to
'that the analysis occurred to the linguist'

(Mulder, 1992, 151)

4. Old English psych verbs

4.1. Data collection

The investigation reported here is based on Old English data from the *York–Toronto–Helsinki Parsed Corpus of Old English Prose* (Taylor et al., 2003) and focuses on instances where a psych verb occurs with more than one syntactic argument (including reflexive verbs, which have one thematic argument), leaving out uses in the passive.

While the goal was to include as many tokens as possible, a number of verbs are not attested in the corpus, possibly being limited to literary use such as *(a/ge–)clian* 'to frighten, excite' or *(on–)egan* 'to fear'. In addition, some verbs were also eliminated due to time constraints given their spelling is ambiguous with non-psych interpretations : *hatian* 'to hate' or 'to get hot'; *freogan* 'to free' or 'to honour'; *gyrnan/geornan* 'to desire' or 'to require'; *ge-/a*-lathian/laethian 'to hate' or 'to invite'. However, this should not undermine the quality of the data given the high number of verbs analysed and the fact that those verbs at first instance did not appear to behave differently from the ones presented here. In total, 1552 tokens of 42 verbs were analysed.

4.2. Nominative Experiencer verbs

The English diachronic syntax literature has included these verbs under the term 'impersonal', even though they are in fact 'personal' constructions, with a nominative Experiencer subject and exhibiting subject–verb agreement. The verbs that fall under this category are in Table 1.

While the Experiencer is always nominative, the second argument, Theme, may vary in case/category. The possible configurations are listed and exemplified below.

4.2.1. Experiencer–NOM and Theme–GEN

(21) *he besargode swiðor his gedwyldes*
he.nom saddened greatly his error.gen
'his error saddened him immensely'

(coaelive,+ALS [Martin]:159.6066)

4.2.2. Experiencer–NOM and Theme–PP

(22) *þa yrsode se casere for his ingange*
then angered the emperor.nom for his entrance

Verb	Translation	No. of tokens
(ge\a–)belgan	to feel anger	4
æfestian	to envy	5
(ge–)andian	to envy	15
(ge–)cweman	to feel pleasure, delight	2
(a\an\on–)drædan	to fear, dread	418
(ge–)fægnian	to rejoice, be glad	96
(a–)forhtian	to be afraid, tremble	72
agrisan	to dread, fear	2
hreowan	to rue, make sorry	4
(be–)hreowsian	to be sorry, repent, lament	81
irsian	to be angry	22
(ge–)lician	to feel pleasure	9
(ge–)lustfullian	to rejoice, be glad	11
(ge–)lystan	to feel pleasure	8
(be–)sargian	to be sad, lament, be sorry for	35
(for–)sceamian	to feel shame, be ashamed	19
ge–swencan	to feel/be disturbed, troubled	1
(an\on–)þracian	to be afraid, fear	4
tintregan	to feel tomented at/with	2
(ge–)tweogan	to hesitate, feel doubt	38
wynsumian	to exult, be joyful	50
Total		899

Table 1: Nominative Experiencer verbs

‘then the emperor got angry for his entrance’

(coaelive,+ALS [Martin]:671.6403)

4.2.3. Experiencer–NOM and Theme–ACC

(23)	*þæt*	*we*	*sculon*	*ure*	*synna*	*behreowsian*
	that	we.nom	should	our	sins.acc	lament

‘that we should regret our sins’

(coaelive,+ALS[AshWed]:37.2719)

This is a fairly common pattern and appears with verbs such as *andian*, *hreowsian*and *sargian*. Such constructions are rarely mentioned in the literature, with the exception of Elmer (1981), given that the focus has normally been on impersonals and these examples have survived into ModE.

4.2.4. Experiencer–NOM and Theme–Clause

I include here those instances where the clausal argument may be a finite clause as in (24) or non–finite.

(24) *þa fægnode Datianus þæt he funde swylcne dry*
then rejoiced Datianus.nom that he found such wizard
'then Datianus rejoiced at finding such a wizard'

(coaelive,+ALS [George]:59.3096)

4.2.5. Experiencer–NOM and Theme–DAT

There is only one instance of this construction in the corpus, pointing to a very idiosyncratic use.

(25) *þa ongon he lustfullian þæs biscopes wordum*
then began he.nom enjoy the bishop's words.dat
'then he begun to enjoy the bishop's words'

(cobede,Bede 2:8.122.32.1168)

4.3. Oblique Experiencer verbs

As in ModE, OE Experiencers may also be non–nominative, appearing marked as either dative or accusative. I will use the term 'oblique' rather than 'object' to remain neutral with regards to their potential subjecthood as it will be discussed in section 6.

Similarly to verbs with nominative Experiencer, the non–Experiencer argument of these verbs may vary in its constituent category and case. The attested combinations are listed below. I will address the issue of what the thematic role assigned to the second argument is in section 5.

4.3.1. Experiencer–DAT/ACC + NOM

(26) *niwe wite abregeð symble þæs mannes mod*
new punishment.nom alarm continually the man's spirit.acc
'the new punishment continually alarms the spirit of the man'

(cogregdC,GD 2 [C]:16.135.16.1633)

4.3.2. Experiencer–DAT/ACC + GEN

(27) *him wlatode þære gewilnunge*
him.dat loathe the desire.gen
'he loathed the desire'

Verb	Translation	Case of Experiencer	No. of tokens
(ge/a–)belgan	to anger, annoy	DAT/ACC	32
(a/ge–)bregan	to frighten, terrify	ACC	7
(ge/mis–)cweman	to (dis)please	DAT/ACC	61
(ge–)egesian	to frighten, terrify	ACC	15
(ge–)eglian	to trouble	DAT	18
aforhtian	to frighten	ACC	1
gegremian	to provoke, irritate	ACC	28
(ge–)hreowan	to cause/feel regret	DAT	16
hreowsian	to pity	ACC	1
(ge–)irsian	to anger	Ambiguous DAT/ACC	3
(ge/mis/of–)lician	to (dis)please	DAT	233
(ge–)lustfullian	to please	DAT/ACC	8
(ge–)lystan	to cause pleasure	DAT/ACC	29
ofþyncan	to regret	DAT	19
(be-)pæcan	to deceive	ACC	27
retan	to comfort	ACC	8
(for–)sceamian	to cause shame	DAT/ACC	59
swencan	to trouble	ACC	26
aþreotan	to displease	Ambiguous DAT/ACC	2
(ge–)tintregan	to torment	ACC	17
(ge–)tweogan	to cause doubt	DAT	41
wlatian	to cause loathe	DAT	3
Total			653

Table 2: Oblique Experiencer verbs

(coaelhom,+AHom 21:89.3130)

4.3.3. Experiencer–DAT/ACC + Clause

Similarly to verbs with nominative Experiencer, the clausal argument may be a finite or an infinitival clause.

(28) *Ne sceamige nanum menn þæt he anum lareowe his gyltas cyðe*
not shame no man.dat that he one master his guilts said
'no man should be ashamed of telling his guilts to a master'
(coaelive,+ALS[Ash Wed]:167.2794)

4.3.4. Experiencer–DAT/ACC + PP

(29) *and swa-ðeah him twynode be his æriste;*
and nevertheless they.dat doubt by his resurrection

‘and nevertheless there is doubt in them about his resurrection’

(cocathom2,+ACHomII,16:162.47.3592)

4.3.5. Experiencer–DAT + ACC

Instances of these constructions are typical cases of impersonal verbs, with the verb selecting a dative Experiencer and an accusative Theme, so the clause has no nominative argument. I found one unambiguous example of it in the corpus, shown below, and three others where it cannot be demonstrated with certainty that the Theme is accusative, with them possibly being nominative. These are one instance of *sceamian* and two cases of *hreowan*,

(30) *him* *gelicade* *hire* *theawas*
him.dat pleasure their virtues
‘to him there was pleasure because of their virtues’

(cochronD,ChronD[Classen-Harm]:1067.35.2283)

4.4. Variation

As well as the case marking variation in verbs with Experiencer object (accusative/dative), many verbs alternate between nominative Experiencer and object Experiencer uses as the reader may have noted. The variation probably reflects the motivations for the change they have undergone over the course of history. Here is a list of the verbs that occur in both types, Nominative and Oblique Experiencer.

Nominative and Oblique Experiencer verbs

Verb	Translation	Tokens
(ge/a–)belgan	to cause/feel annoyance	36
(ge/mis–)cweman	to cause/feel (dis)pleasure	63
(a–)forhtian	to cause/feel fear	73
(ge–)hreowan	to cause/feel regret	20
(be–)hreowsian	to cause/feel pity	82
(ge–)irsian	to cause/feel anger	25
(ge/mis/of–)lician	to cause/feel pleasure	242
(ge–)lustfullian	to cause/feel joy	19
(ge–)lystan	to cause/feel pleasure	37
(for–)sceamian	to cause/feel shame	78
(ge–)swencan	to cause/feel trouble	27
(ge–)tweogan	to cause/feel doubt	79
Total		781

Table 3: Verb alternations

4.5. Reflexives

The occurrence of reflexive psych verbs in OE has not been previously noted in the literature surveyed (mentioned in section 2. Their occurrence is restricted to verbs with nominative Experiencer (or those that alternate between nominative Experiencer and object Experiencer uses), varying between a reflexive and a non–reflexive use as in (31) and (32) respectively. They may also occur with a non–Experiencer argument as in (33). No exclusively reflexive verbs were found.

	Nominative Exp			Oblique Experiencer
	With theme	No theme	Non–RFL	
(a(n)/on–)drædan *'to fear'*	266	18	134	-
(ge/a–)belgan *'to annoy'*	8	16	4	32
forhtian *'to be afraid'*	2	1	69	-
lystan *'to like'*	1	-	7	30
Total	277	35	214	62

Table 4: Reflexive verbs

(31) *he ða gebealh hine*
he.nom then angered himself.acc
'he was then angry'
(cocathom1,+ACHomI,26:395.189.5122)

(32) *ge belgaþ wið me*
you.nom anger with me
'you are angry with me'
(cowsgosp,Jn [WSCp]:7.23.6298)

(33) *hie for þæm hie gebulgon*
they.nom at it themselves.acc angered
'they were angry at it'
(coorosiu,Or 2:8.51.32.990)

Other verbs such as *swencan* and *sceamian* were also found with a reflexive pronoun but these do not seem to be reflexive verbs and were rather simple cases of subject/object co-reference or instances where a reflexive pronoun is used emphatically, as it is the case with the sentence below.

(34) *þonne heo hi sylfe geswnceð in tearum for ðam luste þæs*
when she.nom her self.acc afflicts in tears for the desire the
heofonlican rices
heavenly power
'when she brings herself to tears for the desires of the heavenly empire'
(cogregdC,GDPref and 3 [C]:34.246.10.3483)

5. Analysis

With a description in place, it is now possible to present an account of the syntax of these verbs on the basis of LSE.

Verbs with nominative Experiencer fall straightforwardly into Class I and behave in the same way as they do in ModE. The Experiencer receives nominative case and is the subject, meeting the Experiencer subjecthood requirement of (12). Themes may have their case features checked by [acc] on v, or get their case assigned lexically by V if genitive or dative.

In order to account for instances of PP Themes, the c–selectional features of these verbs must allow for a PP or a DP object, as it is the case in ModE.

(35) John worried about the new project

With regards to reflexives, these constructions are similar to those in Romance languages such as Spanish (where Theme is often optional), and can be acconuted in the same way (see Reinhart, 1997; McGinnis, 1999).

(36)	*Juan*	*se*	*enfadó*	*(con*	*Maria)*
	Juan	himself	angered	(with	Maria)

‘Juan got angry (with Maria)’

As for verbs with Experiencer object, recall that LSE proposes that they should differ depending on the case of the Experiencer. Accordingly, OE accusative Experiencer verbs would fall under Class II, and dative Experiencer ones under Class III.

There has been some discussion in the literature relating to what thematic role of the second argument in these constructions is: while authors such as Visser (1963-73) and Fischer and van der Leek (1983) claim these verbs have a causative interpretation and, therefore, select a Causer, others like Denison (1993) and Allen (1995) have either remained neutral or claimed they are ambiguous between Causer and Theme.[7]. Framing OE psych verbs within LSE should allow for a settling of this issue but it seems that a precise characterisation of which Θ–role is projected may not always be possible in OE.

LSE proposes that Class II may be used statively or eventively, with this aspectual distinction reflecting the thematic role assigned by the verb (Theme or Causer). However, this ambiguity cannot be easily resolved for OE given that the agentivity tests that could detect the existence of a Causer cannot be performed due to the nature of the data available - exactly the same problem that had to be faced in previous research that lead to the disagreement mentinoed in the previous paragraph.

(37)	*&*	*deofles*	*bearn*	*swa*	*swiðlice*	*motan*	*cristene*	*bregean*
	And	devil	son.nom	so/very	strongly	may	christians.acc	terrify

‘and the devil’s son may terrify christians so strongly’[8]

(cowulf,WHom 5:53.206)

Sentence (37) may be interpreted eventively, where the devil’s son terrifies the christians on purpose, say by (appearing in their dreams and) threatening to torture them (if one believes in demonic spirits and took dreams seriously), or statively, if something else caused these people to be scared of the devil's son (a story they heard, for example). However, given that in LSE this ambiguity is restricted to Class II verbs (i.e. those with accusative Experiencers), it should follow that the ambiguity is restricted to verbs with accusative Experiencer but similar problems emerge for verbs with dative Experiencer. Resolving these issues and testing such predictions fall outside the scope of this paper, and it is questionable whether it could be achieved, especially since it has been attempted by notable scholars without coming to a definite answer.

The situation is clearer when the non–Experiencer argument is genitive. If genitive case assignment is inherent and inherent case can only be assigned to internal arguments, genitive DPs are Themes because Causers are selected by v and, thus, are external arguments.

According to LSE, OE dative Experiencers verbs should fall into Class III and select a second argument Theme, and would, thus, be unaccusatives. However, it is not clear if all the non–Experiencer arguments of these verbs are unambiguously Themes. It seems perfectly possible that *mon* 'one, man' in (38) be either a Causer or a Theme.

(38)	*ure*	*yfelnys*	*him*	*oft*	*abelge*
	your	evilness.nom	him.dat	frequently	annoyed

'your evilness annoyed him frequently'

(coaelhom,+AHom2:232.366)

As with OE accusative Experiencers (OE Class II), resolving this potential ambiguity in OE is problematic. If an eventive, causative interpretation for these constructions is indeed possible, the Class II/Class III distinction in OE would be blurred or inexistent, and the only difference between these verbs would be the [acc] or [dat] of the Experiencer. This could be directly related to the subsequent loss of the accusative–dative case distinction, the change undergone by OE psych verbs or a combination of the two. As with Class II verbs, there are serious difficulties in determining whether these verbs have an external argument or not and given this, I will maintain the division proposed in LSE for ease of exposition and classify OE dative Experiencer verbs as OE Class III verbs, leaving open the possibility of Class III–eventives.

The case checking possibilities proceed as follows. The Experiencer is assigned accusative or dative lexically by V as previously mentioned. Non–Experiencer arguments in the nominative have their case features checked structurally by T and, if genitive, they receive lexical case by V, with some mechanism satisfying [nom] on T (e.g. a null expletive). The checking of accusative on non–Experiencers mentioned in section 4.3.5 may be either structural or inherent.

LSE predicts that all object Experiencers are PPs (see section 3) but the corpus did not contain any examples of it. This is a consequence of the productive case morphology of OE which opts to mark the 'locativeness' of Experiencers with case rather than using an overt preposition.

6. Quirky Experiencers in Old English

The theory of psych verbs presented in section 3 proposed a typological division for quirky Experiencers. In this section, potential evidence for the subjecthood of oblique Experiencers is discussed, although it is limited to a restricted subset of psych verbs, specifically those where: i) an oblique Experiencer is used in conjunction with a nominative argument (Theme or Causer); ii) none of the arguments is pronominal given they behave as clitics (see Koopman, 2002).

In order to detect potential cases of quirky subjects, I have collected OE V2 main clauses with non–argumental topics (i.e. adjuncts) where the third clausal constituent is an oblique Experiencer and precedes a nominative DP, and vice–versa. As for subordinate clauses, the only condition was that the Experiencer preceded a nominative DP and vice–versa.

(39)

ðurh	*andan*	*bepæhte*	*se*	*deofol*	*þone*	*frumsceapenan*	*mann*
through	malice	deceived	the	devil.nom	the	first-created	man.acc

'the devil deceived the first created man through enmity'

(cocathom1,+ACHom I,39:523.90.7877)

(40)

ða	*gelicode*	*Gode*	*þeos*	*ben.*
then	pleased	God.dat	this	prayer.nom

'then this prayer pleased God'

(cocathom2,+ACHom II,45:336.30.7538)

(41)

ðætte	*hiera*	*Dryhten*	*licige*	*ðæm*	*folce*
that	their	lord.nom	pleased	the	people.dat

'that their lord pleased the people'[9]

(cocura,CP:19.147.5.994)

(42)

gif	*þam*	*gifran*	*angemetlic*	*spræc*	*ne*	*eglde*
If	the	greedy.dat	eloquent	speech.nom	not	afflicted

'if the greedy are not afflicted by loquacity'

(cocura,CP:43.309.2.2069)

While there were examples where an accusative Experiencers appeared after the nominative DP, no accusatives preceded the nominative argument.

	Matrix	**Subordinate**
NOM before ACC/DAT	2	11[10]
DAT before NOM	3	9
Total	3	9

Table 5: Word order in clauses with an oblique Experiencer and a nominative DP

Although the numbers are small, the frequency of Experiencers preceding the nominative DP in these constructions is striking, occurring just as often as 'canonical' nominative first constructions. If they are subjects, this seems to be in parallel with Mulder's (1992) point that such word order can be the default case (see section 3.4).

But could these preposed Experiencers be subjects? All three main clause tokens found have an initial *þa*, which have been shown to be unambiguous cases where the finite verb (Vf) moves to C (see SP). The question that must be asked now is whether these dative DPs are in fact in subject position or are VP–internal, with Spec,TP filled by a null expletive.

If we adopt Williams's (2000) proposal that an empty expletive in declarative clauses is only licensed if c–commanded by Vf, the examples in question could be analysed as containing a null expletive as Vf is in C, where it c–commands Spec,TP. Thus, the dative Experiencer subject would be ruled out.

However, Williams (2000) also argues that constructions with locative inversion in OE (and Middle English) do not have a silent expletive which would leave Spec,TP empty. As it is well known, subject position cannot normally be left unnocupied given the EPP requirement. Under this perspective, what could satisfy it? If we recall that in LSE Experiencers are mental locations and as such they must be locally related to T (the LF-subjecthood requirement; see section 3, it may be the case that the dative Experiencer is what satisfies the EPP, as a result of locative inversion, or "Experiencer–inversion", which raises the DP to Spec,TP. Although this cannot be demonstrated, this analysis would suggest that these dative Experiencers are quirky subjects.

What about subordinate clauses? In all 9 clauses the dative Experiencer precedes Vf and could potentially be subjects. However, an analysis of their position runs into the problem of whether they are subjects or topics. There is no consensus as to whether topicalisation in OE was possible in both main and subordinate clauses (symmetrical V2) as SP argues, or it was restricted to main clauses (asymmetrical V2) as claimed by VK.[11] Topicalisation in main clauses allows clear identification of what is its subject and what is its topic but the unclear status of topicalisation in embedded clauses means that there is no way to distinguish whether a clause-initial Experiencer DP takes such position to satisfy the topic requirement or as a result of EPP–triggered movement to subject position. In (42) for example, the dative Experiencer could be either a subject or a topic, with a head final Infl (AgrS and T). As a result of this, there is also uncertainty with respect to what is the position for embedded subjects – Spec,vP for SP and Spec,TP for VK.

Consequently, it is impossible to evaluate van Kemenade's (1997) claim that preposed Experiencers in embedded clauses are not topics and appear because 'impersonal verbs' are unaccusative. In addition, the discussion in section 5 showed that an analysis on the basis of unaccusativity cannot be tested as it is not clear when a verb with Experiencer object is unaccusative, so predictions cannot be made. However, from the perspective of LSE and the argument presented here, it is the 'locativeness' of Experiencers and their quirkiness that leads to their fronting, rather than the lack of an external argument leaving subject position unnoccupied and available to non–nominatives.

As an analysis of OE quirky Experiencers in embedded clauses cannot be developed whilst remaining neutral with regards to embedded V2, I will consider it under the assumption that topicalisation is restricted to main clauses and Spec,TP is the position for subjects as proposed by VK, given it predicts that fronted datives are not topics.

From this perspective, the 9 embedded DAT–first clauses should be analysed as having a quirky Experiencer. The null expletive licensing condition mentioned above rules out the existence of a silent expletive in these clauses as Vf does not c–command Spec,TP: in embedded clauses, Vf moves to T but not to C, as it is already occupied by a complementiser, so Vf remains in T. This is true for at least 7 of the 9 clauses found, given that two are complements of a bridge verb which involve CP–recursion, where Vf may be in C (see Iatridou and Kroch, 1993). A null expletive analysis would also be ruled out if one adopted the same locative/experiencer–inversion argument based on Williams (2000) mentioned above. If these clauses do not have a silent expletive, the initial position of these dative Experiencers would be a result of movement to Spec,TP to satisfy the EPP–feature of T, making them quirky subjects. This would answer the question raised in van Kemenade (1997, 334) where she leaves open the possibility of these dative DPs being in subject position or inside the VP.

Although the evidence is not conclusive, the data discussed here gives further support to Allen's (1995) claim for the subjecthood of OE non–nominative Experiencers. She argues that OE accusative and dative Experiencers can be subjects since they appear to control deletion of the subject of a coordinated clause as demonstrated in the example below.

(43)

ac	*gode*	*ne*	*licode*	*na*	*heora*	*geleafleast*	*ac*	*asende*	*him*	*to*	*fyr*
but	god.dat	not	liked	not	their	faithlessness.nom	but	sent	them	to	fire
of	*heofonum*										
of	heaven										

'but their faithlessness did not please God, but (he) sent them fire from heaven'

Relating this to LSE's typology in (13), OE would pattern with Icelandic, a Scandinavian language, in that it allows both accusative and dative quirky subjects, rather than a Continental Germanic language like Dutch which allows only dative subjects.

7. Conclusion

This paper presented the range of constructions that appear with OE psych verbs and how they fit into a general syntactic account of these verbs, explaining their syntactic properties (case marking, word category, θ–marking). It was shown that verbs with an oblique Experiencer are aspectually ambiguous and given the difficulties in resolving it due to the nature of the data it is not possible to give a precise characterisation of whether they are unaccusatives and what non–Experiencer θ–role they assign. In the last section, some evidence was presented suggesting that OE dative Experiencers may be quirky subjects which, combined with Allen (1995), would place OE alongside Scandinavian languages in a typology of quirky subjects.

Notes

1. The argument presented here is also consistent with a more articulate left periphery and/or an AgrO projection.

2. Adger (2003) does not mention inherent case assignment.

3. There is no consensus with regards to the subject position in embedded clauses (Spec,TP in VK; Spec,VP in SP).

4. 'CS' indicates that the bracketed string refers to a conceptual structure and 'INCH' is the abbreviation of inchoative.

5. LSE follows Pesetsky (1995) in dividing Belletti and Rizzi's (1988) characterisation of Themes into Causer, Target of Emotion and Subject Matter. For simplicity, I will use Theme to refer to Target of Emotion and Subject Matter, given the distinction is irrelevant here.

6. Square brackets are used to refer to (un)interpretable features.

7. Although Allen (1995) argues for the Causer–Theme distinction in at least some verbs

8. Lionarons (2000).

9. van Kemenade (1997).

10. Two tokens of this configuration were excluded given the nominative argument was the clitic–like *man* 'one, man'.

11. More recently, it has also been suggested that OE was not a V2 but a V3 language (Kroch, 2009).

References

Adger, David. 2003. *Core Syntax*. Oxford: Oxford University Press.

Allen, Cynthia L. 1995. *Case Marking and Reanalysis*. Oxford: Oxford University Press.

Anagnostopoulou, Elena. 1999. On Experiencers. In *Studies in Greek Syntax*, ed. Artemis Alexiadou, Geoffrey C. Horrocks, and Melita Stavrou. Dordrecht: Kluwer.

Arad, Maya. 1998. VP-structure and the Syntax-Lexicon Interface. Doctoral Dissertation, University College London.

Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: Chicago University Press.

Barðal, Jóhanna. 1999. The Dual Nature of Icelandic Psych-Verbs. *Working papers in Scandinavian Syntax* 64:79–101.

Belletti, Adriana, and Luigi Rizzi. 1988. Psych-Verbs and θ-theory. *Natural Language and Linguistic Theory* 6:291–352.

Bresnan, Joan. 1994. Locative Inversion and the Architecture of Universal Grammar. *Language* 70:72–131.

Chomsky, N. 2001. Derivation by phase. In *Ken Hale: a life in language*, ed. K.L. Hale and M.J. Kenstowicz. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, ed. David Michels Roger Martin and Juan Uriagereka. Cambridge, MA: MIT Press.

Denison, David. 1993. *English Historical Syntax*. New York: Longman.

Elmer, W. 1981. *Diachronic grammar: the history of Old and Middle English subjectless constructions*. Niemeyer.

Fischer, Olga, Ans van Kemenade, Willem F. Koopman, and Wim van der Wurff. 2000. *The Syntax of Early English*. Cambridge: Cambridge University Press.

Fischer, Olga, and Frederike C. van der Leek. 1983. The demise of the Old English impersonal construction. *Journal of Linguistics* 19:337–68.

van der Gaaf, Willem. 1904. *The transition from the impersonal to the personal construction in Middle English*. Heidelberg: Winter.

Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, MA: MIT Press.

Haeberli, Eric. 2000. Adjuncts and the Syntax of Subjects in Old and Middle English. In *Diachronic Syntax: Models and Mechanisms*, ed. George Tsoulas Susan Pintzuk and Anthony Warner. Oxford: Oxford University Press.

Iatridou, Sabine, and Anthony Kroch. 1993. The licensing of CP-recursion and its relevance to the Germanic verb-second phenomenon. *Working Papers in Scandinavian Syntax* 50:1–24.

Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, MA: MIT Press.

Jespersen, Otto. 1984. *Progress in Language*. London: Allen & Unwin.

van Kemenade, Ans. 1987. *Syntactic Case and Morphological Case in the History of English*. Dordrecht: Foris.

van Kemenade, Ans. 1997. V2 and embedded topicalisation in Old and Middle English. In *Parameters of Morphosyntactic Change*, ed. Ans van Kemenade and Nigel Vincent. Cambridge: Cambridge University Press.

Koopman, Willem F. 2002. Another look at Clitics in Old English. *Transactions of the Philological Society* 95:73–93.

Kroch, Anthony. 2009. Was Old English a verb-second language? Talk given at the University of York, UK.

Landau, Idan. 2005. *The Locative Syntax Of Experiencers*. Ms. Unpublished.

Landau, Idan. 2009. *The Locative Syntax of Experiencers*. Cambridge, MA; MIT Press.

Lionarons, J. T. 2000. Wulfstan's Eschatological Homilies. Retrieved online at http://webpages.ursinus.edu/jlionarons/wulfstan/wulfstan.html.

McGinnis, Martha. 1999. Reflexive clitics and the specifiers of vP. *MIT Working Papers in Linguistics* 35:137–160.

Mulder, Rene. 1992. The aspectual nature of syntactic complementation. Doctoral Dissertation, Holland Institute of Generative Linguistics, The Netherlands.

Pesetsky, David. 1995. *Zero Syntax*. Cambridge, MA: MIT Press.

Pintzuk, Susan. 1993. Verb seconding in Old English: Verb movment to Infl. *The Linguistic Review* 10:5–35.

Pintzuk, Susan. 1995. Verb-seconding in Old English. In *Historical Linguistics 1993 (Current Issues in Linguistic Theory, 124)*, ed. H. Andersen. Amsterdam/Philadelphia: John Benjamins.

Pintzuk, Susan. 1999. *Phrase Structures in Competition: Variation and Change in Old English Word Order*. New York: Garland.

Reinhart, Tanya. 1997. Syntactic effects of lexical operations: Reflexives and unaccusatives. *UiL OTS Working Papers in Linguistics* .

Taylor, Ann, Anthony Warner, Susan Pintzuk, and Frank Beths. 2003. *The York Toronto Helsinki Parsed Corpus of Old English Prose*. Available through the Oxford Text Archive.

Visser, Fredericus Th. 1963-73. *An Historical Syntax of the English Language*. Leiden: Brill.

Williams, Alexander. 2000. Null subjects in Middle English existentials. In *Diachronic Syntax: Models and Mechanisms*, ed. George Tsoulas Susan Pintzuk and Anthony Warner. Oxford: Oxford University Press.

SEQUENTIALLY DETERMINED FUNCTION OF PITCH CONTOURS: THE CASE OF ENGLISH NEWS RECEIPTS

MARIANNA KAIMAKI

Abstract

This paper examines the pitch characteristics of a responsive turn, the news receipt, in order to explore the relationship between the prosodic design of a turn and its impact on the way conversation evolves. The results suggest that at this place in interactional structure the prosodic design of the news is dependent on whether the turn involves an explicit positive or negative lexical assessment (valenced news receipt) or not (non-valenced news receipt). When a non-valenced news receipt is employed the option of employing a falling or a rising pitch contour is open to the producer of the turn. When a valenced news receipt is employed there is only the option of a falling pitch contour. Non-valenced news receipts employing falling and rising pitch contours were found to receive the same range of interactional treatments suggesting that rising and falling pitch contours are in free variation – they can alternate without affecting the 'meaning' and further development of conversation at least at next turn position. This finding is particularly important as it suggests that at some places in interactional structure the function and meaning of pitch contours might not be the same as at some other place in structure.

1. Introduction

CA research has shown that participants systematically display, in the placement and design of their own talk, an understanding of each others' talk and of the actions which that talk implements. One of the actions performed by participants in interaction is the imparting and receiving of news. The news can be good or bad and it makes relevant a response, a news receipt, by the co-participant. The way people respond to a news telling is obviously important. It can determine the further development of conversation as there are sensitive issues of preferred/dispreferred kinds of responses to news. For example, some news receipts have explicit valency (i.e. oh that's great) while some others do not (i.e. oh really). Using a valenced instead of a non-valenced news receipt is a choice participants in conversation have and their choice may influence the subsequent development of talk. In this paper sequential and phonetic analysis of news tellings are made in an attempt to determine whether there are lexical, intonational or other cues that participants employ/orient to in imparting/receiving news that have an effect in the further development of talk and consequently the way the news are being treated. The results are used to ground analyses of phonetic variation and phonological organisation in the observable behaviours and reactions of the participants themselves.

The function of intonation to date has been associated mainly with grammatical (e.g. question/statement intonation, information focus) or 'paralinguistic' (e.g. assertiveness, submissiveness, uncertainty, anger) aspects of meaning (Gussenhoven 2002, 2004; Ladd 1996 Cruttenden 1997). These remain particularly elusive as linguists have not yet provided statements of linguistic contrasts for them in a similar fashion they managed to do so with 'segmental' material of speech. Researchers on the interface of Conversation Analysis and phonetics have chosen to look at 'bundles' of phonetic features (Local 2007) instead and have associated clusters of phonetic characteristics with managing conversational actions such as turn-taking, repair initiation or collaboratively completing a turn (e.g. Auer 1996; Couper-Kuhlen 1992, 1996, 2001a,b; Curl

2005; Curl, Local, and Walker 2004; Ford and Thompson 1996; Fox 2001; French and Local 1986; Local 1986, 1992, 1996, 2000, 2004; Local, Kelly, and Wells 1986; Local and Walker 2006; Ogden 2006; Selting 1996, 1998; Szczepek-Reed 2004; Walker 2004; Wells and Local 1993; Wells and Macfarlane 1998; Wells and Peppé 1996). These researchers view conversation as the locus for linguistic analysis and base their claims on the behaviour of participants in conversation. The way a participant in conversation designs their talk affects the way their co-participant receipts their turn which in turn has consequences for the development of conversation. The 'cues' for differential treatment of turns can be a conglomeration of phonetic features which are context specific. This paper adopts the methodology employed by those researchers but introduces the notion of free variation of phonetic features in managing actions in conversation. The conversational sequence examined is the 'news informings' sequence where free variation in the choice of pitch contours is observed.

News informings are used by conversational participants to impart some allegedly not previously known information to a co-participant. The action of telling some news is not done in a haphazard way. It is instead orderly and follows a sequential pattern previously analysed by Schegloff (1988) and Maynard (2003). A sequential pattern has been proposed for the News Delivery Sequence (NDS) (Maynard 2003: 95, see also Maynard 1997): A typical News Delivery Sequence is composed of four turns preceded by an optional turn of a preannouncement of the news by the news teller or an enquiry by the recipient of the news. After a news announcement a response to the news is made relevant which can be of three kinds: a news receipt, a newsmark or a standardised *oh*-prefaced assessment. This 3-way distinction is based on functional grounds. The production of an announcement response of the *oh really*, *oh lovely* etc kind discourages any elaboration on the news on the part of the participant imparting the news. Newsmarks on the other hand encourage elaboration on the news just imparted. The third type of announcement response – standardized *oh*-prefaced assessments – is equivocal in terms of function and can either encourage or discourage elaboration of news. After the announcement response turn, a turn elaborating on the news just imparted may be relevant depending on the kind of announcement response produced by the receiver of news. Then the news sequence is usually ended by an assessment of the news.

Maynards characterisation of the News Delivery Sequence appears to overdetermine its organisation. Not all sequences involving the telling of news have the sequential design announcement of news – announcement response – elaboration of news – assessment of news. In particular, a number of researchers have observed that news informings responded to *by oh really* have the following sequential design: '(1) News announcement, (2) *oh really* (3) reconfirmation and (4) assessment' (Jefferson 1981). Heritage (1984: 324) has shown that certain kinds of news receipts (standalone *oh*) can curtail news telling sequences. Local (1996) also demonstrates that news receipts of the form 'oh+assessment' routinely occur 'at the termination of a topic/news-informing' (1996: 189).

Although news receipts have attracted a good deal of interactional analytic attention, there has not been a comparable amount of research on their phonetic design. However, Local (1996) provides a phonetic description of some *oh* and *oh*-prefaced news receipts and Freese and Maynard provide a description of the phonetic design of news receipts in terms of valence (Freese and Maynard 1998:195). One thing missing is a consideration of the design of a broader set of news receipts and an examination of the extent to which their phonetic design is consequential for how they are treated and the subsequent development of the talk.

The paper is organised as follows. Section 2 describes the data considered for this study. Section 3 presents the phonetic characteristics associated with non-valenced news receipts. Section 4 presents the phonetic characteristics associated with valenced news receipts. Section 5 explores the sequential organisation of news receipts. Finally section 6 brings together the results.

2. Data

The data used in this paper consists of naturally occurring conversational speech drawn from English and American telephone calls between family and friends. The data represents all the instances of a subset of valenced and non-valenced news receipts in the following corpora: Callhome (transcribed parts only), Holt, NB, Kamunsky, MTRAC, SBL, Rahman, Heritage and HG. The total duration of the audio recordings amounts to some 77 hours. Table 1 shows the number of instances and percentages of the kinds of news receipts found in the data.

Table 1: Number of instances/percentages of English news receipts

English news receipts			
Kind	**Structure**	**N**	**%**
Valenced	Oh+more	118	62%
Non-valenced	*oh really*	73	38%
Total		**191**	**100%**

Oh may be followed by *really* or *oh* by *more talk*. Under the *oh+more* category there are four subcategories: *oh+adj*, *oh that's+adj*, *oh how+adj* and *oh very+adj* which will be analysed below.

3. Phonetic characteristics of non-valenced news receipts

Overall, there are 73 measurable *oh really* tokens produced as single intonational phrases in the data examined. In terms of pitch contour 21 of these are produced with a final falling contour while 52 are produced with a final rising contour. This does not seem to depend on the variety of English used, whether British or American. All tokens are done as single intonational phrases. The most prominent part of these tokens is *really*.

There are 21 instances of *oh really* tokens produced with falling pitch contours in the dataset. Of these only 18 proved amenable to robust phonetic analysis (3 of the tokens were either in overlap in a mono channel sound file or the poor quality of the recording did not allow for any acoustic measurements). The canonical overall shape of the falling tokens of *oh really* is as follows: there is a rise associated with *oh* (mean rise = 7st, max = 16.9st, min = 0.8st) which is followed by a falling contour associated with *really* (mean fall = 10.9st, max = 21.8st, min=1.8st).

Instrumental analysis of the tokens show that the start of the fall – i.e. the highest F_0 point before the onset of the fall – is aligned in two ways with the articulatory material: early or late. Figures 1[1] and 2 show pitch traces which give a visual impression of the two different kinds of alignment of the falling tokens of *oh really*.

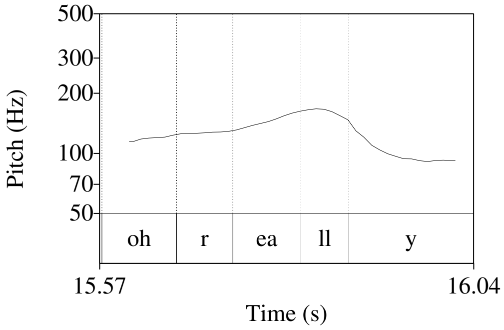


Figure 1: F_0 trace with articulatory annotation illustrating pitch peak occurring during the lateral (en5273-274).

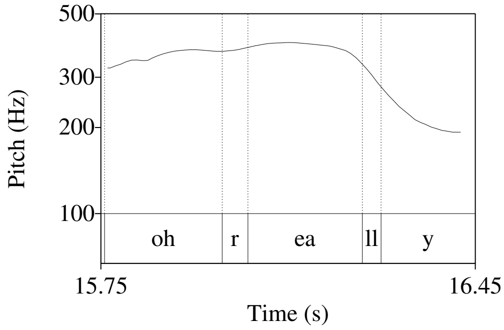


Figure 2: F_0 trace with articulatory annotation illustrating pitch peak occurring during the vocoid before the lateral (en4571-294).

The pitch trace in figure 1 shows a token of *oh really* in which the highest F_0 point is late-aligned and occurs during the lateral. This is the most frequent pattern observed in the data with 11 tokens of *oh really* having this kind of alignment. Figure 2 shows a pitch trace of an *oh really* news receipt where the main pitch peak occurs during the vocoid before the lateral. Early alignment of the highest F_0 was found in only 7 of falling *oh really* tokens.

The second kind of pitch contour employed in news receipts done with *oh really* is a rising pitch excursion. There are 52 cases in our data displaying rising pitch. The general pattern in the tokens is for the pitch to be level (or display a small fall) from *oh* to the beginning of *really* and then to rise during *really*. The range of the rise may vary, and rises with the greatest excursion end high in the speaker's range (mean rise = 7.5st, max = 16.7st, min = 1st, SD = 4). Whatever the extent of the range, the rise always ends higher than the starting point of any fall in pitch associated with *oh*. As with the falling tokens of *oh really*, there is variability in the alignment of pitch with articulatory material. Over half the rising tokens have an early aligned turning point of the rise (during the *oh* portion (7%) or the onset (51%) of *really*). For the remainder the turning point is aligned later with first syllable vocoid (23%) or the lateral of *really* (19%). Figures 3 and 4 show representative cases of early- and late- aligned *oh really* tokens.

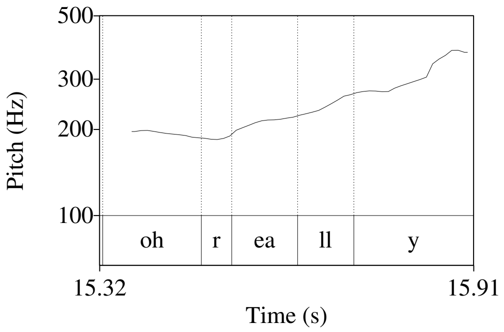


Figure 3: Pitch trace of rising early-aligned *oh really* with a TP during the rhotic (en4665-1451)

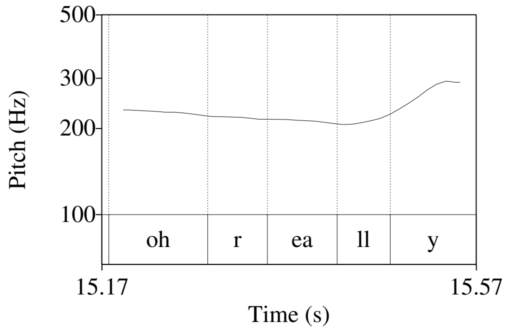


Figure 4: Pitch trace of rising late-aligned *oh really* with a TP during the lateral of *really* (en4576-615)

4. Phonetic characteristics of valenced news receipts

There are 118 measurable instances of valenced news receipts of the *oh+more* structure (i.e. *oh+adj*, *oh that's+adj*, *oh how+adj*, *oh very+adj*) in the data examined. In terms of pitch contour 114 of them are produced with a final falling contour while 4 are produced with a final rising contour[2]. This does not seem to depend on the variety of English used, whether British or American.

4.1. OH+ADJ

The first subcategory of the *oh+more* valanced news receipts examined is the *oh+adj* category (82 tokens in total) in which the adjectives encountered were: *good* (61), *great* (7), *dear* (3), *nice* (3), *cool* (2), *terrible* (1), *marvelous* (1), *lovely* (1), *brilliant* (1), *neat* (1) and *perfect* (1). *Oh+adj* news receipts are routinely done with an overall fall (mean range = 11st, max = 23st, min = 3st , SD = 5). In the majority (56) of these the falling contour is associated with the adjective. In 22 of these the starting point of the fall occurs during the *oh* (there is no statistical difference in pitch range between *oh+adj* turns with falling contours where the fall co-occurs with *oh* or with the adjective).

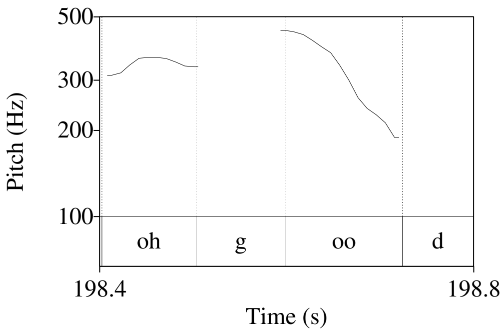


Figure 5: Late alignment *oh good* (en4576-198)

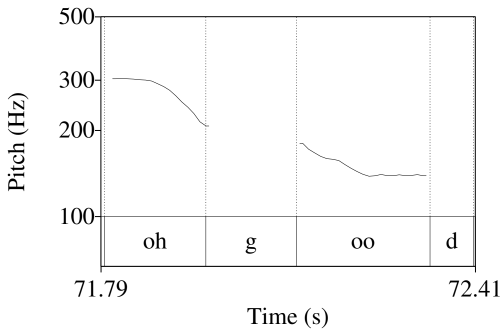


Figure 6: Early alignment *oh good* (Holt.M.88.1.2-71)

Figure 5 shows the most common pattern of pitch alignment of the *oh+adj* news receipts. As can be seen, there is an initial step-up from *oh* to the beginning of *good* which is where the highest F_0 point occurs and where the beginning of the fall starts. The falling pitch contour co-occurs with the (stressed) adjective and the end of the fall is lower than the pitch where the step up was made from. Figure 6 shows the second most common pitch pattern of the *oh+adj* news receipts in which the highest F_0 point occurs during the (stressed) *oh*. As can be seen, the start of the fall occurs during the production of *oh* and continues through the adjective (*good*).

4.2. OH THAT'S+ADJ

The second subcategory of *oh+more* valenced news receipts is *oh that's+adj* (27 tokens in total). The 'adj' component of which is occupied by the following adjectives: *good* (11), *great* (6), *wonderful* (3), *nice* (3), *neat* (1), *swell* (1), *marvellous* (1) and *a pity* (1). As with the majority of the *oh+adj* cases, *oh that's+adj* are done with a final falling contour (mean range = 10st, max = 19st, min = 3st, SD = 6). The majority (17 out of 27) are produced with a falling contour where the starting point of the fall and major accent peak is during 'that's. The other 10 instances are done with the fall done over the adjective (i.e. 'oh THAT's x' versus 'oh that's X')[3]. Figures 7 and 8 and show representative cases of each pitch contour and alignment.

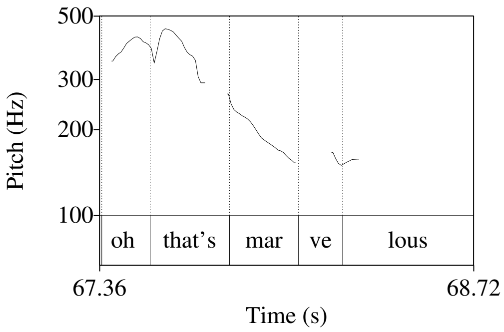


Figure 7: Early alignment of turning point of fall (with 'that') in *oh that's+adj* (Holt.M.88.1.2-67)

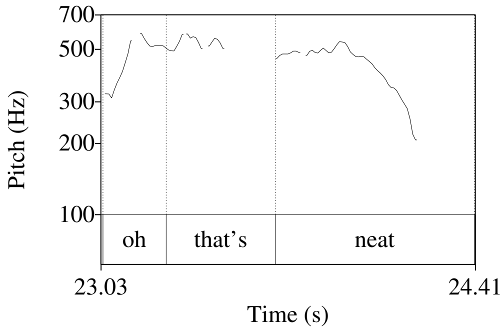


Figure 8: Late alignment of turning point of fall (with 'neat') in *oh that's+adj* (KamunskyIII-23)

4.3. OH HOW+ADJ

The third subcategory of *oh+more* valenced news receipts is *oh how+adj*. In this subcategory the adjectives encountered are *nice* (1), *lovely* (3), *cool* (1) and *gross* (1). Unlike the *oh+adj*, *oh how+adj* news receipts are only done with an overall fall. 3 out of 6 cases are done with a fall (or a step down) the highest F_0 point of which is on *how* or *oh* and the remaining 3 instances are done with a fall which takes place over the adjective. Figures 9 and 10 show examples of the two different alignments of the falls.

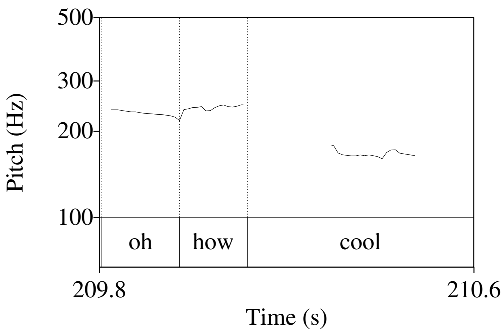


Figure 9: Early alignment *oh how+adj* (en6348-209)

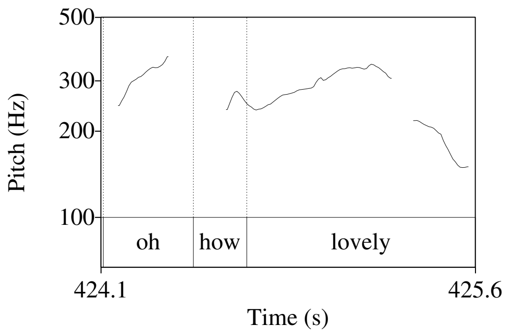


Figure 10: Late alignment *oh how+adj* (Holt.88.2.4-261)

Figure 9 shows a valenced news receipt (*oh how cool*) produced with an overall falling pitch contour with early alignment. The highest F_0 point is aligned with *how* followed by a step down to a relatively level F_0 coincident with the adjective. Figure 10 shows a pitch trace of the valenced news receipt *oh how lovely*. The highest F_0 point of this news receipt occurs during the adjectival part of the receipt.

4.3.1. OH VERY+ADJ

The last subcategory of the *oh+more* valenced news receipts is the *oh very+adj* group. It should be noted that unlike the previous categories where both positively and negatively valenced adjectives were encountered, *oh very adj* news receipts only displayed positively valenced adjectives.

The adjectives in this subcategory are: *good* (2) and *nice* (1). The subcategory of *oh very+adj* is similar to the *oh how+adj* subcategory in that it displays only falling pitch contours. In terms of alignment, the start of the fall is on *nice* in the case of *oh very nice* while for the *oh very good* cases it varies between the first syllable of *very* and the second syllable of *very*. Examples of these different kinds of pitch alignments are shown in figures 11, 12 and 13.

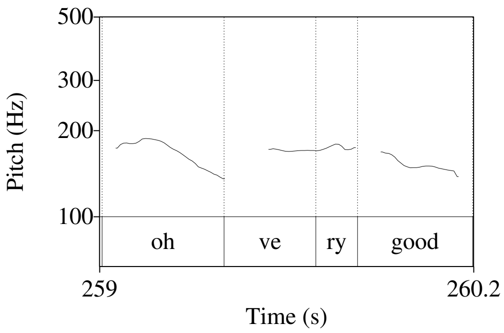


Figure 11: Second syllable alignment *oh very+adj* (en4941-258)

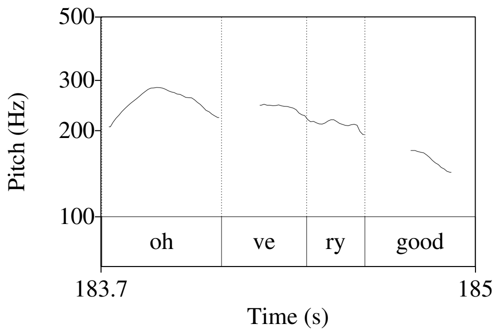


Figure 12: First syllable alignment *oh very+adj* (Rahman.II-183)

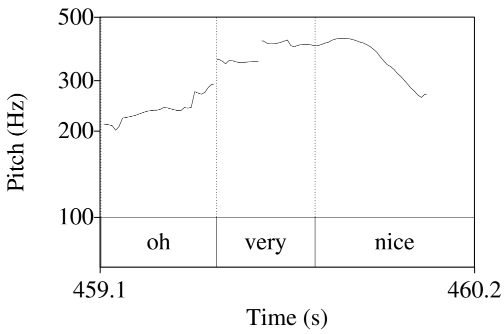


Figure 13: Late alignment of *oh very+adj* (en4623-459)

Figure 11 shows a valenced news receipt *oh very good* in which the highest F_0 point of the fall is aligned with the second syllable of *very*. Figure 12 shows the pitch trace of another instance of the valenced news receipt *oh very good*. Here, the highest F_0 pitch point of the fall coincides with the production of the first syllable of *very*. Finally, figure 13 shows a pitch trace of a valenced news receipt of the structure *oh very nice* where the highest F_0 point of the fall (which is also the highest F_0 point of the utterance) is aligned with the beginning of the adjectival part of the news receipt.

5. Interactional/sequential organisation

The previous sections showed that there is variability in the phonetic make-up in terms of pitch contour shape within the same category (i.e. falling *oh really* and rising *oh really*). There were also differences observed in the alignment of pitch with the articulatory material of falling and rising *oh really*. These will also be discussed later in subsection 5.1. This section explores whether the sequential organisation of the different kinds of news receipts varies systematically. In doing this we ask whether there are any differences and if so whether they can be accounted for by looking at the prosodic or structural design of news receipts. First, we examine the sequential structure of news receipts done with rising or falling tokens of *oh really*.

5.1. Non-valenced news receipts: oh really

News sequences with rising tokens of *oh really* have one of the structures shown in Table 2[4].

Table 2: Rising *oh really* structures

Rising *oh really*

		Structure	**N**	**%**
Structure 1	Speaker A:	Informing		
	Speaker B:	*oh really*		
	Speaker A:	confirmation		
Total			**10**	**24%**
Structure 2	Speaker A:	Informing		
	Speaker B:	*oh really*		
	Speaker A:	confirmation+elaboration of news		
Total			**14**	**33%**
Structure 3	Speaker A:	Informing		
	Speaker B:	*oh really*		
	Speaker A:	continuation of informing		
Total			**10**	**24%**
Structure 4	Speaker A:	Informing		
	Speaker B:	*oh really*+more talk		
Total			**8**	**19%**
Grand Total			**42**	**100%**

As can be seen in Table 2, *oh really* news receipts produced with a final rising contour are treated in four different ways in next-turn position. I treat the combination of informing, receipt and treatment as discriminable structures:

- in Structure 1 Speaker A produces an informing which is receipted by Speaker B with a rising pitched *oh really*. In next turn position Speaker A issues a confirmation of the news;
- in Structure 2 Speaker A produces an informing which is receipted by Speaker B with a rising pitched *oh really*. After Speaker A confirms the news, they do not give up the turn but continue with an elaboration of the news imparted in their preceding turn;

- in Structure 3 Speaker A produces an informing which is receipted by Speaker B with a rising pitched *oh really*. Speaker A then continues with an elaboration of the news they imparted in their preceding turn without giving any confirmation of the news;
- in Structure 4 Speaker A produces an informing which is receipted by Speaker B who does not relinquish their turn but continues to produce more talk.

Sequences of news informings involving falling *oh really* tokens have one of the structures reported in table 3.

Table 3: Falling *oh really* structures

Falling *oh really*

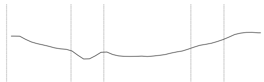
		Structure	N	%
Structure 1	Speaker A:	Informing		
	Speaker B:	*oh really*		
	Speaker A:	confirmation		
Total			**3**	**14%**
Structure 2	Speaker A:	Informing		
	Speaker B:	*oh really*		
	Speaker A:	confirmation+elaboration of news		
Total			**4**	**19%**
Structure 3	Speaker A:	Informing		
	Speaker B:	*oh really*		
	Speaker A:	continuation of informing		
Total			**5**	**24%**
Structure 4	Speaker A:	Informing		
	Speaker B:	*oh really*+more talk		
Total			**9**	**43%**
Grand Total			**21**	**100%**

As Table 3 indicates, the next-turn types following *oh really* tokens produced with falling contours are the same as those reported for sequences with rising *oh really* tokens. These results suggest that though there is variation it does not systematically differentiate the two types of pitch pattern: there does not appear to be a systematic difference in the kinds of next turn uptake between the falling and rising tokens. However, there is a difference in the frequency of occurrence of these next-turn types following falling or rising *oh really* receipts. While a confirmation and an elaboration on the news is the most frequent sequential structure (Structure 2) for *oh really* news receipts done with a rising contour (14/42), more talk by the speaker producing the news receipt (Structure 4) is the most frequent sequential organisation for news informings receipted with falling tokens of *oh really* (9/21). Although there appears to be differences between the distribution of rising and falling tokens of *oh really* across the structures, the relatively small number of falling tokens means that quantitative interpretation is not straightforward.

Fragment (1) shows a canonical instance of a rising *oh really* token with a Structure 1 organisation. Speaker A produces some information at lines 1 and 2 about her boyfriend's parents. Speaker B receipts this information at line 4 with a rising *oh really*. Immediately after, at line 5 speaker A confirms the truthfulness of the news with *yeah*. In next turn position B embarks on an enquiry – which is related to the news informing at lines 1-2 – about whether A's boyfriend can speak French. This is followed by A's subsequent response at line 7.

(1) en 4157-723

1 A: which is good (0.2) and uh (0.5) his dad works at a bank
2 INF → (0.2) his mother is French
3 (0.8)

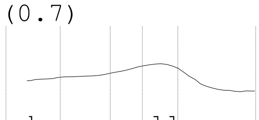


4 B: oh r ea ll[y]
5 REC → A: [y]eah [((she))]
6 B: [does he] speak French
7 A: yes he does, fluently

Fragment (2) shows the sequential organisation of an informing where a token of *oh really* with a falling pitch contour is employed.

(2) en 5273-274

1 A: but yeah ((you know)) my mom finally got email
2 INF → (0.7)



3 B: oh r ea ll y
4 REC → A: yeah
5 B: what's her email address
6 (0.5)
7 A: oh I don't remember off hand

It can be seen that the sequential structure of the informing is similar to the one reported before for the rising *oh really* tokens. There is a production of an informing at line 1, which is receipted by a non-valenced news receipt in the form of a falling pitched *oh really*. The news receipt then is followed by a confirmation ('yeah') of the news by the producer of the informing at line 4. We can observe that there are other structural similarities between these two fragments. For instance, in the turn immediately after the turn containing the 'yeah' confirmation the news recipients embark on talk which solicits further talk derived from the telling. At fragment (1) speaker B at line 6 produces an enquiry – which is derived from the news informing – about whether A's boyfriend speaks French. This is followed by a response at line 7 by speaker A. Similarly, in fragment (2) speaker B immediately after the 'yeah' confirmation embarks on an enquiry – which is again derived from the news telling at line 1 – about the e-mail address of A's mum. This is followed by a 0.5 secs pause and a subsequent response by A at line 7.

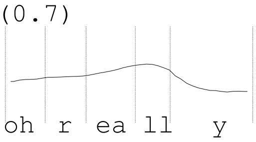
F_0 alignment

Both falling and rising tokens of *oh really* were found with two kinds of alignment of the F_0 pitch peak: late and early. By late alignment is meant that the F_0 high peak of the fall is aligned with the lateral or the last vocoid of *oh really* and by early alignment that the F_0 high peak of the fall is aligned with any articulatory material present before the lateral. Examination of conversational fragments with the differently aligned *oh really* tokens shows that the interactional environment and uptake of differently aligned *oh really* tokens is similar in all cases. There does not seem to be a direct mapping between alignment of the F_0 highest point of the fall and any major differences in sequential structure. Consider the following fragments.

Fragment (2), reproduced here as Fragment (3), illustrates a sequence where a late-aligned *oh really* is employed.

(3) en 5273-274

```
 1          B:   you does your email I mean are you getting the emails I
 2               think you are I mean [you respo]nded o[kay]
 3          A:                        [ y e a h ]      [mhm] Usually
 4               whenever I get a message I respond in some way or other
 5               huh huh huh huh huh huh huh huh huh
 6          A:   (clearing throat) sorry you know me
 7               (0.3)
 8          A:   .hhhh
 9               (0.5)
10          A:   but yeah ((you know)) my mom finally got email
11 INF →         (0.7)
```

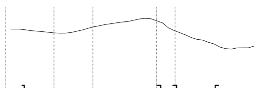


```
12          B:   oh r ea ll  y
13 REC →    A:   yeah
14          B:   what's her email address
15               (0.5)
16          A:   oh I don't remember off hand
```

It can be seen that the uptake of the news receipt conforms to the structures reported for the falling *oh really* non-valenced news receipts in table 3. The non-valenced news receipt is followed by a confirmation ('yeah') in next turn position. Fragment (4) shows an informing sequence where an early-aligned *oh really* is employed.

(4) en 5713-549

```
 1          B:   how's the Aztec doing
 2               (0.2)
 3          A:   uh- loves it- first day out in the weather just (it) was
 4 INF →         running around all playful like almost playful with the-
 5               the dog and the goat will walk together now
 6               (0.7)
```



```
 7          B:   oh r  ea ll [y   ]
 8 REC →    A:               [yeah] they'll walk they'll walk side by
 9               side kind of ignoring each other but if either one stops
10               the other one will wait
```

Speaker A informs B about the well-being of a farm goat and its relationship with his dog. Speaker B receipts this information with a non-valenced news receipt (*oh really*) in which the highest F_0 point of the fall is early-aligned occurring during the production of the vocoid. Speaker A confirms the news at line 8 with 'yeah' and continues to provide more information about the behaviour of the animals. Examination of these two data fragments reveals that despite the fact the differently aligned falling tokens of *oh really* get the same treatment i.e. a confirmation in the form of *yeah* as an immediate response they don't have the same kinds of sequential development after the confirmation.

Data analysed so far suggests that there may be a difference in the options available to co-participants when using one alignment over the other. It seems that if an early aligned *oh really* is employed the co-participant has the option of doing a confirmation followed by more talk but does not have the option of just doing a standalone confirmation.

5.2. Valenced news receipts: oh+adj, oh that's adj, oh how+adj and oh very+adj

Examination of valenced news receipts in the data suggests that their sequential organisation is similar to the patterns documented for the non-valenced news receipts. The uptake after the production of the news receipt is mainly done by a continuation of the informing or the production of *yeah*. However, closer examination reveals that there are subtle but potentially important differences in the sequential structure of valenced news receipts. These differences will be discussed in section 6 below. Tables 4[5], 5 and 6 show the kinds of patterns encountered in the data for the *oh+adj* (design structure A), *oh that's adj* (design structure B) and *oh how+adj*, *oh very+adj* (design structure C) news receipts:

Table 4: Valenced news receipt: design structure A

Valenced news receipt: *oh+adj*

		Structure	**N**	**% out of GT**
Structure 1	Speaker A:	Informing		
	Speaker B:	*oh+adj*		
	Speaker A:	continuation of informing		
Total			**46**	**61%**
Structure 2	Speaker A:	Informing		
	Speaker B:	*oh+adj*		
	Speaker A:	ratification (+assessment)		
Total			**16**	**21%**
Structure 3	Speaker A:	Informing		
	Speaker B:	*oh+adj*		
	Speaker A:	other (assessment, continuers)		
Total			**4**	**5%**
Structure 4	Speaker A:	Informing		
	Speaker B:	*oh+adj*+more talk		
Total			**10**	**13%**
Grand Total			**76**	**100%**

Table 5: Valenced news receipt: design structure B

Valenced news receipts: *oh that's adj*

		Structure	N	% out of GT
Structure 1	Speaker A:	Informing		
	Speaker B:	*oh that's adj*		
	Speaker A:	continuation of informing		
Total			**13**	**48%**
Structure 2	Speaker A:	Informing		
	Speaker B:	*oh that's adj*		
	Speaker A:	ratification		
Total			**9**	**33%**
Structure 3	Speaker A:	Informing		
	Speaker B:	*oh that's adj*		
	Speaker A:	assessment, more talk enquiry		
Total			**5**	**19%**
Grand Total			**27**	**100%**

Table 6: Valenced news receipt: design structure C

Valenced news receipts: *oh how+adj*, *oh very+adj*

		Structure	N	% out of GT
Structure 1	Speaker A:	Informing		
	Speaker B:	*oh how+adj*	5	63%
	Speaker A:	continuation of informing		
Structure 2	Speaker A:	Informing		
	Speaker B:	*oh how+adj*	1	11%
	Speaker A:	ratification		
Total			**6**	**67%**
Structure 1	Speaker A:	Informing		
	Speaker B:	*oh very+adj*	1	11%
	Speaker A:	continuation of informing		
Structure 2	Speaker A:	Informing		
	Speaker B:	*oh very+adj*	1	11%
	Speaker A:	ratification		
Structure 3	Speaker A:	Informing		
	Speaker B:	*oh very+adj*	1	11%
	Speaker A:	same assessment		
Total			**3**	**33%**
Grand Total			**9**	**100%**

As can be seen, the structural and lexical make-up of a news receipt does not seem to be responsible for differences in the further development of talk. At next turn position after the production of the news receipt all design structures (A, B and C) have the option of 'continuation of informing'. Ratifications are also encountered in next turn position after the production of the news receipt in all design structures A an B and C. Fragments (5) to (8) show cases where different structural make ups are treated in the same way in next turn position. Fragments (5) and (6) show examples of an *oh+adj* and a *oh that's adj* news receipt which are both followed by ratification of their appropriateness by the producer of the informing. Fragments (7) and (8) show examples of news receipts with different adjectives which are followed by continuation of the informing irrespectively.

In fragment (5) Joy is telling Lesley about the painful consequences of a recent medical procedure her husband has had. At line 1 Joy produces a strong assessment of the amount of pain her husband experiences which Lesley responds to with a valenced news receipt (*oh dear*) followed by laughter. In overlap with Lesley's laughter Joy ratifies the correctness of Lesley's valenced news receipt ('yes') and goes on to produce a further assessment. Similarly, in fragment (6), B informs A about the working days she has left in a job she hates and how happy she is she gets to finish earlier than her contract (lines 1-4). At line 5, A receipts this information with a positively valenced receipt (*oh that's great*) and B in the next turn ratifies A's valenced assessment as being appropriate for the informing she produced.

(5) Holt May 88 1.2

```
1          Joy:  but uh oh it must be terribly painful I can hear him
2  INF →          when he's in the bathroom going
3                 oh hu-ow hu-[oh
4          Les:               [oh dear[ hheh hha huh ha  ].h h h h h
5  REC →   Joy:                        [ yes  it must  be ] terrible
```

(6) en 4157-304

```
1          B:   I hate my job so much oh my God I can't wait like next
2  INF →        .hhhhhh well my last day is Thursday but um
3          A:   [((oh c-))]
4          B:   [I'm  only] scheduled till Tuesday I'm so happy
5  INF →   A:   oh that's great
6  REC →   B:   yeah [.h  ]
7          A:        [well] do you need the extra time though
8          B:   yeah
9          A:   because it's coming up fast
```

In fragment (7) B informs A of sister Molia's and sister Katerina's welfare which A receipts at line 4 with a positively valenced news receipt of the *oh that's adj* structure. B then at line 5 continues to give more information on the same topic – notice the use of 'and' – which is also receipted with a positively valenced receipt of the *oh that's adj* type by A. At line 8 B continues giving more information on the topic which A receipts with a continuer ('mhm'). Fragment (8) shows an informing sequence where two news receipts of the *oh+adj* format are being used at lines 11 and 15. As can be seen these are also followed by more talk on the same topic by the producer of the informing (lines 12 – 13, 16 – 17).

(7) en5208 – 528

```
          B:  sister Molia is good .h
INF →     A:  [( * * ) ]
          B:  [sister K]aterina is very good
INF →     A:  oh [that's great]
REC →     B:     [.hh   and  s]ister um (0.2) .h
INF →         Molia is standing here smiling at me ha [ ha     .hhh  ]
          A:                                          [oh that's good]
REC →     B:  and I told her that you were going to call
          A:  mhm
```

(8) en6456 – 801

```
          A:  okay sh- she was probably the educational one this is the
              religious h
          B:  oh okay
              (0.2)
          A:  um and she had just done an article with ministers on the
INF →         call
              (0.4)
          B:  uh-h[uh ]
          A:      [she] referred to that a little bit and that was in
INF →         last Sunday
          B:    [           oh   good          ]
REC →     A:  .h[hhh so this should it's going to] be quite interesting
              to see what she does with this
              (0.5)
          B:  oh great
REC →     A:  we'll save you a copy or send you a copy we'll probably
              send you a copy
```

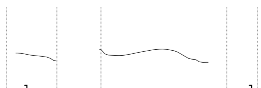
The above fragments show that the structural and lexical make up of news receipts do not appear to determine what happens in next turn position or to affect the development of the subsequent talk. Irrespective of structure *oh+adj* vs *oh that's adj* or lexical make up (*good* vs *great*) valenced news receipts seem to permit the same range of possibilities in their uptake. However, there is a noticeable absence of some kinds of turns after the production of a news receipt in some design structures within the valenced news receipt category. This suggests that structural designs may be treated differently by co-participants. For example, when news is receipted with a structure C design (see Table 6 on page 13) there is only ever continuation of the news telling in next turn position. News receipts of the *oh that's adj* are never followed by more talk by the same speaker while news receipts of the *oh+adj* design are regularly found to do so.

F_0 *alignment*

As was shown in section 4 valenced news receipts are found with both early and late alignments. Early-aligned *oh+adj* cases have the F_0 pitch peak aligned with the *oh* while late-aligned *oh+adj* cases have the highest F_0 pitch point occurring during the adjective. Interactional analysis indicates that there is no difference in the uptake between the early- and late-aligned *oh+adj* news receipts. Fragments (9) and (10) show examples of differently aligned valenced news receipts which exhibit similar interactional treatments.

(9) en4705-1554

```
 1            B:   that’s when Peggy called me that night
 2            A:   oh (0.2) [that’s right]
 3            B:            [Peggy   Doug]herty
 4            A:   that’s right
 5                 (0.2)
 6            B:   a[nd  ] and uh was telling me a little
 7            A:    [yeah]
 8            B:   and I felt so bad I hadn’t written some of it
 9                 [but] I did get letters off to them
10  INF →     A:   [oh ]
11            B:   [.h notes] off to them anyway
12  INF →     A:   [    oh  ]
13            B:   [.hhhh] with Don
14  INF →     A:   [ oh  ]
```

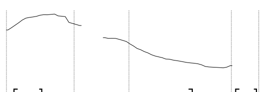


```
15            A:   oh  g   oo    d
16  REC →     B:   so yeah and that was that was so good the mail down here
17                 is terrible do you know last week I went to the ((Fe
18                 Allegria)) main office.
```

In fragment (9) speaker B produces an informing at lines 9 -13 to which speaker A responds with a late-aligned valenced news receipt – *oh good*. At line 16 after an initial ‘so’ which possibly projects an upshot of the telling speaker B ratifies speaker A’s receipt as being appropriate for the news she just imparted (‘yeah’) produces an upgraded assessment ‘so good’ and then continues with more talk related to her informing.

(10) Rahman I -163

```
 1            Jen:  (nyeo) it was a little bit saucy hhh [heh  heh ]
 2            Ver:                                       [oh was it]
 3            Ver:  ehh[hehh heh h]ah ha[h hah ]
 4            Jen:      [y(h)ehheh ]     [.ah .h]hh and it was i (.) you know
 5  INF →          it was a right good mur[der[right good]thrill[er]
 6            Ver:                             [y  [y  e  s   ]      [m ]
 7            Jen:  [mm         ]
```



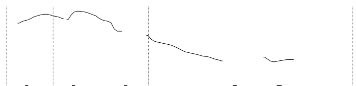
```
 8            Ver:  [oh   g    o]o:[d
 9  REC →     Jen:                    [.hh yes .h and it (ended) you know in
10                 sort of an it ended in a great big bang
11                 ehhh he[h  h n I  jum]ped out of=
12            Ver:        [oh    huh    ]
13            Jen:  =the e seat I [jumped
```

In fragment (10) Jenny informs Vera at lines 1 - 5 about a film she has watched recently. Vera responds to the news with a positively valenced news receipt, *oh good*, produced with an early peak alignment. In next turn position Jenny ratifies Vera's news receipt as being appropriate for the news she imparted and then continues with more talk about her informing. We can notice that in both fragments (9) and (10) following the confirmation token the continuation of the informing is designed with *and*.

Early- and late-aligned valenced news receipts of the *oh that's adj* structure were also found in the data. In early aligned tokens of *oh that's adj* the highest F_0 point co-occurs with *that's*, while in late aligned tokens the highest F_0 point occurs during the adjective. Early-aligned and late-aligned tokens are treated in the same way by co-participants; both may get continuation of informings or ratifications in next turn position. Fragments (11) and (12) exemplify cases where both early- and late-aligned *oh that's adj* structures get the same uptake in next turn.

(11) Holt.M88.1-2 59

```
1          Joy:  =[and um .p I went to pick him up (.) Mister Williams
2                came in an said .hhh you'll be glad to know I checked him
3   INF →        and .hhh ( ) it's as (0.3) clear as a whistle
4                he[said
```



```
5          Les:   [.hhhh Oh that's  mar[velous.]
6   REC →  Joy:                        [  n o  ]: no problems at all he
7                said aren't you pleased I said yes
8          Les:  oh good
9          Joy:  so he said we'll see him again next year
```

In fragment (11) Joy produces an informing at lines 1- 4 about what the doctor said regarding the condition of her husband. Leslie receipts the news, at line 5, with a positively valenced receipt of the *oh that's adj* structure in which the F_0 pitch point of the fall is early-aligned. At next turn position Joy continues and elaborates her informing.

In fragment (12) speaker A is providing her co-participant with a positively framed account of some of the activities she's been engaged in since her move to a new locale.

(12) en4595–451

```
          A:  I did- the first month that I was here I took a class in
INF →         digital music [um which is like e]lectronic
          B:                [   oh      cool   ]
          A:  you know like [music] editing
INF →     B:                [ mhm ]
          A:  and stuff like tha[t ]
INF →
```

```
          B:                     [oh]that's   gre[at   ]
REC →     A:                                    [and c]ause I've been
              w- you know working with sound and I'm interested in
              working more with sound and stuff and so that was really
              good we [there was a concert at] the end of it
          B:          [   that's     great   ]
          A:  and I felt (0.6 ) pretty good about my contribution to
              that like .h[hh i]t wasn't really exactly what I would
          B:              [.hh ]
              want but I felt like .hhhhhhhh l[ike  I   ]
          B:                                  [oh that's] great
          A:  yeah so that was g- that was good that was something that
              you know like y- being here gave me access to
```

At line 7 speaker receipts this information with a late-aligned positively valenced ‘oh that’s great’. At next turn position, speaker A continues with her informing. Examination of the data indicates that different F_0 alignment, within structural subcategories, does not systematically condition differences in the organisation of the ensuing talk. Early-aligned and late-aligned *oh that’s adj* and *oh+adj* news receipts can have the same uptake in next turn position.

5.3. *Some conclusions*

There are two noticeable differences in the sequential organisation of the valenced news receipts as compared to the non-valenced news receipts. The first is that valenced news receipts may get ratifications by the producer of the news, while non-valenced news receipts get confirmations of the informing by the producer of the news. The second is that while in valenced news receipts there is an option for the speaker who imparted the news to produce an assessment after the news receipt there is no such option in the non-valenced news receipts.

Results of the analysis also showed that differences in the prosodic phonetic make up of news receipts do not seem to affect the sequential structure of the news informing sequence. For example news receipts with falling and rising pitch contours may have the same uptake and are treated in the same way by co-participants. However, alignment of the pitch contour does seem to have an effect on the sequential organisation of news informing sequences, as some of the early- and late-aligned tokens displayed different sequential designs.

The next section brings together the findings for valenced and non-valenced news receipts and discusses the theoretical implications of our findings.

6. Discussion

Results showed that pitch contours vary within each group and across the two groups of valenced and non-valenced news receipts in terms of alignment and dynamic pitch movement (falling/rising). Interactional organisation is fairly uniform within individual groups. However, comparison of the interactional organisation of the two groups (valenced – non-valenced) revealed a consequential difference with respect to the choice of actions available in the two kinds of news receipts. When an informing is receipted with a valenced news receipt (*oh that's+adj oh+adj*) by a co-participant one of the choices the news producer has in next turn position is to ratify the news he/she imparted. However, when an informing is receipted with a non-valenced news receipt (e.g. *oh really*) the possible action choices available to the news teller do not involve ratifying the news but confirming the news he/she imparted.

Interactional analysis of valenced and non-valenced news receipts shows that the relationship between pitch variation and sequential structure is not a straightforward one. More specifically, tokens of rising and falling *oh really* have each a set of identical choices of next-turn structures which are:

1. a confirmation;
2. a confirmation+elaboration of news;
3. continuation of informing;
4. more talk by the speaker producing the news receipt.

This suggests that rising and falling pitch contours are in free variation in this interactional context. The choice of one over the other does not result in different uptakes by coparticipants at next turn position. The same actions can be done after the production of either a rising or a falling news receipt. So, variation between rising and falling pitch contours does not seem to have any observable sequential consequences immediately after the production of the news receipt. There might be interactional evidence for rising and falling news receipts to be treated differently further on in the development of conversation. This will require a more extensive analysis of the sequences examined here.

The frequency with which participants employ uptakes 1 to 4, however, differs between the rising and falling tokens of *oh really*. Rising tokens of *oh really* are mostly followed by a confirmation (with or without elaboration) of the news by the next speaker (57% of the rising tokens) while falling ones are mostly followed by more talk by the same speaker (43% of the falling tokens). This finding suggests that a falling pitch contour can have the function of a 'continuation rise' as both rising and falling tokens of *oh really* may be followed by more talk. Furthermore, it is the falling tokens of *oh really* rather than rising tokens which are more frequently followed by more talk. These results suggest the need for an alternative account for the function of falling and rising pitch events in everyday interaction. In valenced news receipts the employment of a falling or (very rarely) a rising pitch contour does not seem to have any effects on the sequential development of the news informing sequence.

Early or late alignment of the F_0 peak of the falling pitch contours does not appear to have any major effect on the sequential structure of either the valenced or the non-valenced news receipts examined. However, findings – although tentative – showed that *there are some differences in the sequential structure between news receipts with different pitch alignments that are worthy of note*. It was observed that an early- aligned *oh really* news receipt provides the co-participant with the option of doing a confirmation followed by more talk but not with the option of doing a confirmation without more talk following (as is the case with late aligned tokens of *oh really*). On the other hand, early and late aligned valenced news receipts with falling pitch do not seem to affect the subsequent development of conversation in any way. Both early and late aligned tokens of *oh that's+adj* and *oh+adj* are always followed by either continuation of informings or ratifications in next turn positions.

The sequential organisation of news receipts was found to be more variable than those reported by Local (1996) and Maynard (2003). For instance, it is not always the case that non-valenced *oh really* news receipts have the sequential design: (1) news announcement, (2) *oh really*, (3) reconfirmation, (4) assessment. Analysis shows that there are other options available for co-participants in position (3) such as continuation of informing, confirmation and elaboration of the news or more talk by the producer of *oh really*. Maynard's (2003) claims about the different function of news receipts and newsmarks were not borne out either. According to Maynard when news receipts are employed they 'show a retrospective orientation... while acting to deter development of the news' (Maynard 2003: 100) while newsmarks are used to 'promote its (i.e. the news') development and elaboration' (Maynard 2003: 100). One of the announcement responses which Maynard recognises as a news receipt is *oh really*. According to Maynard's assumptions then news receipts done with *oh really* would act as a deterrent for the producer of the news to offer more information on the news he/she just imparted. However, analysis of the data in this paper reveals that one of the possible actions a co-participant does after an *oh really* news receipt is to continue with the elaboration of the news he/she has just imparted. The results also provide robust confirmation of Local's (1996) observation that *oh really* receipts may be done with either falling or rising final contours.

Valenced news receipts are overwhelmingly produced with an overall falling pitch contour (114/118). This finding robustly confirms the more limited observations made by Local (1996). This is of some interest, in that if there is a constraint such that only a falling contour may co-occur with valenced news receipts, the linguistic 'value' of the falling contour cannot be the same as that of a falling contour which can occur at the same place in structure as a rising contour – it cannot be doing the same linguistic 'work'. This echoes the view of the Firthian prosodic analysts (Firth 1957b; Sprigg 1957). For instance, Sprigg (1957: 264) writes 'Of particular relevance ... are: (i) the concept of *system*, the closed set of commutables in which the value of each of the commuting prosodic or phonematic terms is a function of the number of terms in the system, with the result that no term in a two-term system, for example, can be identified with any term in a system comprising a number of terms other than two (though it may be compared with it)'. Thus there is little to be gained linguistically (phonologically) by equating the falling contour of *oh lovely*, for instance, with the falling contours of an *oh really* token. One of the implications of this is that a polysystemic rather than a monosystemic approach might be more felicitous in accounting for this variability. Valenced and non-valenced news receipts could be viewed as two different structures with different systems of contrast operating within them.

The system of choice of falling/rising pitch contours, however, does not seem to operate in the same way in the two groups. In the non-valenced news receipts the choice of a fall over a rise is 21 (falling) versus 52 (rising) while in the valenced news receipts it is 114 (falling) over 4 (rising). It seems then that non-valenced news receipts favour rising pitch contours and valenced pitch contours favour falling pitch contours. It was noted earlier that there were a small number of tokens of valenced news receipts produced with a final rising contour. These receipts are all instances of the *oh good* structure. The choice of rising pitch contour then in the valenced news receipts seems to be associated with a lexical item. The use of a rising pitch contour with an *oh good* news receipt does not seem to have any effects on the interactional structure of the news sequence. The 4 fragments where *oh good* is employed with a rising pitch contour are reproduced below as fragments (13), (14), (15) and (16).

(13) Holt.1.1

```
                  (0.6)
            Les:  I've got Katharine on them too now
INF →       Mum:  oh good nh hnh hnh huh .hh yeah I'll let you know
REC →           h[nh
            Les:  [yes .h[h
            Mum:         [actually they made me run .h
                  (0.3)
            Les:  .hh oh well u I think you'll get used to that
            Mum:  yes
```

(14) Holt.SO.88.II.13b

```
            A:   are you all going cause I wouldn't like to go on my own
            B:   there's lots of us going
            A:   ye[s yes]
            B:     [ uhm ] erm (0.3) the Morrisons and the Waringses and
INF →            Tommy Barnes and his wife and all sort of peop[le]
            A:                                                  [oh] good
REC →       B:   got about eighty
            A:   .pt .hh oh lovely
```

(15) RahmanI

```
            Jen:  =[I was just eh ringing up to say I'll be coming down in
INF →           a moment
                  (.)
            Ver:  oh g- good
REC →             (.)
            Ver:  good
            Jen:  yah
```

(16) RahmanIb

```
            Ver:  [yes yes she applied she rang up on the Monday morning
INF →           you[know
            Jen:    [m:mm oh good=
REC →       Jen:  =[mm
            Ver:  =[and she's got the application forms
```

The sequential organisation observed in fragments (13) to (16) is:

1. fragment (13) – A: news informing, B: oh good + more talk A: confirmation;
2. fragment (14) – A: news informing, B: oh good, A: continuation of informing;
3. fragment (15) – A: news informing, B: oh good+more talk, A: ratification;
4. fragment (16) – A: news informing, B: oh good, A: continuation of informing.

All of these structures are also observed in *oh good* news receipts with falling pitch contours. In light of the present data then we cannot provide a motivated interactional account for rising *oh good* news receipts[6].

These results have implications for an empirically grounded analysis of falling and rising pitch contours in English and indicate the need for an approach to intonational contrast which takes full cognisance of the sequences such receipts occur in and the actions that they implement. Contrast is a central concept in linguistic theory. However, what does or does not constitute a phonological contrast is handled rather differently in different phonological frameworks. Differences between phonological approaches arise mainly from whether, and how, different levels of linguistic analysis are permitted to interact and what sorts of things can be taken into account as phonological 'context'. For instance, Firthian Prosodic phonology (Firth 1957b,c) takes grammatical, syntactic and lexical considerations into account from the outset. Unlike many styles of phonological analysis, Firthian analyses are polysystemic: they recognise different systems of phonological contrast operating at different places in structure (e.g. at word/syllable initial position versus word/syllable final position; in grammatical versus lexical items (Firth 1957a; Sprigg 1957). The results reported in this paper suggest that intonational systems might be better accounted for by a polysystemic approach with respect to interactional organisation. That is, rather than viewing pitch contrasts in terms of a single system potentially available across a whole language, it is possible to think of place in interactional sequence as par t of the complex of levels to be taken into account when specifying systems of contrast. The result of this would be that pitch variation could be analysed in terms of many systems operating at different places in interactional structure. This would allow us to recognise that pitch could have different kinds of contrastive potential when doing different kinds of interactional work (which may be reflected in the precise detail of its phonetic shape).

Notes

1. The Y-axis in all the figures presented here is labelled Pitch (Hz). No theoretical assumptions are entailed by this usage which is the default output for the 'draw pitch' function in Praat.
2. These 4 cases will be analysed in the discussion section.
3. The small numbers of instances do not permit statistical comparison, but impressionistically there does not seem to be noticeable differences in the pitch range of falling contours associated with different accentual placement or with the different adjectives.
4. The total does not add up to match the total of Table 1 as 10 of the tokens could not be classified under any of the 4 structures proposed.
5. The total does not add up to the total given for *oh+adj* cases in subsection 4.1 because 6 of the cases could not be classified under any of the 4 structures proposed.
6. These rising tokens produced by different speakers do not appear to be idiosyncratic in that at least some of these speakers also produced *oh good* tokens with a canonical falling contour.

References

Auer, P., 1996. On the prosody and syntax of turn continuation. In: Couper-Kuhlen, E., Selting, M. (Eds.), Prosody in Conversation. Cambridge University Press, Cambridge, pp. 57–94.

Couper-Kuhlen, E., 1992. Contextualizing discourse: The prosody of interactive repair. In: Auer, P., di Luzio, A. (Eds.), The Contextualization of Language. John Benjamins, Amsterdam, pp. 337–364.

Couper-Kuhlen, E., 1996. The prosody of repetition: On quoting and mimicry. In: Couper-Kuhlen, E., Selting, M. (Eds.), Prosody in Conversation: Interactional Studies. Cambridge University Press, Cambridge, pp. 366–405.

Couper-Kuhlen, E., 2001a. Constructing reason-for-the-call turns in everyday telephone conversation. InLiSt (25).

Couper-Kuhlen, E., 2001b. Interactional prosody: High onsets in reason-for-the-call turns. Language in Society 30, 29–53.

Cruttenden, A., 1997. Intonation, 2nd Edition. Cambridge University Press, Cambridge.

Curl, T. S., 2005. Practices in Other-Initiated Repair Resolution: The Phonetic Differentiation of ‘Repetitions’. Discource Processes (39), 1–43.

Curl, T. S., Local, J., Walker, G., 2004. Repetition and the prosody-pragmatics interface. York Papers in Linguistics Series 2 1, 29–63.

Firth, J., 1957a. Papers in Linguistics, 1934–1951. Oxford University Press, London.

Firth, J., 1957b. A synopsis of linguistic theory, 1930–1955. In: Studies in Linguistic Analysis. Basil Blackwell, Oxford, pp. 1–32.

Firth, J., 1957c. The technique of semantics. In: Papers in Linguistics, 1934–1951. Oxford University Press, London, pp. 7–33.

Ford, C. E., Thompson, S. A., 1996. Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. In: Ochs, E., Schegloff, E. A., Thompson, S. A. (Eds.), Interaction and Grammar. Cambridge University Press, Cambridge, pp. 134–184.

Fox, B. A., 2001. An exploration of prosody and turn projection in English conversation. In: Selting, M., Couper-Kuhlen, E. (Eds.), Studies in Interactional Linguistics. John Benjamins, Amsterdam, pp. 287–315.

Freese, J., Maynard, D. W., 1998. Prosodic features of bad news and good news in conversation. Language in Society 27, 195–219.

French, P., Local, J., 1986. Prosodic features and the management of turn interruptions. In: Johns-Lewis, C. (Ed.), Intonation in Discourse. Croom Helm, London, pp. 157–180.

Gussenhoven, C., 2002. Intonation and interpretation: phonetics and phonology. In: Bel, B., Marlien, I. (Eds.), Proceedings of the Speech Prosody Conference. pp. 47–57.

Gussenhoven, C., 2004. The phonology of tone and intonation. Cambridge University Press, Cambridge.

Heritage, J., 1984. A change-of-state token and aspects of its sequential placement. In: Atkinson, J. M., Heritage, J. (Eds.), Structures of Social Action: Studies in Conversation Analysis. Cambridge University Press, Cambridge, pp. 299–345.

Jefferson, G., 1981. The abominable ‘Ne?’: a working paper exploring the phenomenon of post-response pursuit of response. Occasional paper no. 6, University of Manchester, Department of Sociology.

Ladd, D. R., 1996. Intonational Phonology. Cambridge University Press, Cambridge.

Local, J., 1986. Patterns and problems in a study of Tyneside intonation. In: Johns-Lewis, C. (Ed.), Intonation in Discourse. Croom Helm, London, pp. 181–198.

Local, J., 1992. Continuing and restarting. In: Auer, P., di Luzio, A. (Eds.), The Contextualization of Language. John Benjamins, Amsterdam, pp. 273–296.

Local, J., 1996. Some aspects of news receipts in everyday conversation. In: Couper-Kuhlen, E., Selting, M. (Eds.), Prosody in Conversation. Cambridge University Press, Cambridge, pp. 177–230.

Local, J., 2000. Phonetic construction of collaborative completions, presented at the EuroConference on Interactional Linguistics, Spa, Belgium.

Local, J., 2004. On the interactional and phonetic design of collaborative completions. In: Hardcastle, W., Beck, J. M. (Eds.), A Figure of Speech: a Festschrift for John Laver. Lawrence Erlbaum, New York, to appear.

Local, J., 2007. Phonetic detail and the organisation of talk-in-interaction. In: XVIth International Congress of Phonetic Sciences, Saarbrucken, Germany. pp. 1–10.

Local, J., Kelly, J., Wells, B., 1986. Towards a phonology of conversation: Turn-taking in Tyneside English. Journal of Linguistics 22, 411–437.

Local, J., Walker, G., 2006. Methodological imperatives for investigating the phonetic organization and phonological structures of spontaneous speech. Phonetica 62, 120–130.

Maynard, D. W., 1997. The news delivery sequence: Bad news and good news in conversational interaction. Research on Language and Social Interaction 30 (2), 93–130.

Maynard, D. W., 2003. Bad News, Good News Conversational Order in Everyday Talk and Clinical Settings. The University of Chicago Press.

Ogden, R., 2006. Phonetics and social action in agreements and disagreements. Journal of Pragmatics 38, 1752–1775.

Schegloff, E. A., 1988. On an actual virtual servo-mechanism for guessing bad news: A single case conjecture. Social Problems 35 (4), 442–457.

Selting, M., 1996. Prosody as an activity-type distinctive cue in conversation: the case of so-called ‘astonished’ questions in repair initiation. In: Couper-Kuhlen, E., Selting, M. (Eds.), Prosody in Conversation: Interactional Studies. Cambridge University Press, Cambridge, pp. 231 – 270.

Selting, M., 1998. TCUs and TRPs: The construction of units in conversational talk. InLiSt 4.

Sprigg, R., 1957. Junction in spoken Burmese. In: Studies in Linguistic Analysis. Basil Blackwell, Oxford, pp. 104–138.

Szczepek-Reed, B., 2004. Turn-final intonation in English. In: Couper-Kuhlen, E., Ford, C. E. (Eds.), Sound Patterns in Interaction. John Benjamins, Amsterdam, pp. 97–117.

Walker, G., 2004. The phonetic design of turn endings, beginnings, and continuations in conversation. Ph.D. thesis, University of York.

Wells, B., Local, J., 1993. The sense of an ending: A case of prosodic delay. Clinical Linguistics and Phonetics 7, 59–73.

Wells, B., Macfarlane, S., 1998. Prosody as an interactional resource: Turn-projection and overlap. Language and Speech 41 (3–4), 265–294.

Wells, B., Peppé, S., 1996. Ending up in Ulster: Prosody and turn-taking in English dialects. In: Couper-Kuhlen, E., Selting, M. (Eds.), Prosody in Conversation: Interactional Studies. Cambridge University Press, Cambridge, pp. 101–130.

Marianna Kaimaki
Department of Language and Linguistic Science
University of York
email: mk502@york.ac.uk

VARIATION AND SIMILARITY IN THE PHONOLOGICAL DEVELOPMENT OF FRENCH DIZYGOTIC TWINS: PHONOLOGICAL BOOTSTRAPPING TOWARDS SEGMENTAL LEARNING?

CATHERINE E. SMITH

Abstract

The few existing studies comparing the language acquisition of twins have shown contrasting findings with regard to phonological development (Bloch (1921), Leonard et al (1980)). Focussing on the implementation of templates – preferred output patterns implemented to facilitate production – this study compares the phonological acquisitions of a pair of twins, as systematisation of the output through templates appears to set the framework for an important shift in language development. Parallel templates are discussed alongside the observation of a coordinated shift towards segmental learning despite an initially dissimilar approach to language learning. Evidence from this data suggests that a form of phonological bootstrapping is taking place between these infants, which leads to further questions regarding the language acquisition of twins in the longer-term.

1. Introduction

Common universals can be found in the phonological acquisition of different children acquiring different languages; Smith (1975) identified four universal tendencies of phonological systematization: consonant harmony, cluster reduction, systematic simplification and grammatical simplification. Examples of these phenomena can be found in many longitudinal studies of individual children, such as Macken's study of Si, acquiring Spanish (1979), Smith's study of Amahl, acquiring English (1973), and Vihman's study of Raivo, acquiring Estonian (1981). In all of these studies, phonological patterns, or 'templates', are implemented to facilitate production in the output, generalizing common features found in the infants' early word forms and causing an infant to adapt words to fit a familiar output pattern (Vihman and Vihman (in press)). Words are either 'selected' or 'adapted' in line with an infant's preferred output patterns; words which already fit the structure of an infant's preferred pattern are 'selected' for production and are thus produced accurately, while target words which do not match an infant's preferred output pattern are 'adapted', changing the structure of the word to match a preferred output pattern. Adapted words highlight an overriding pattern force of a specific template, allowing an infant to produce more words, forfeiting output quality for output quantity (Smith (2010)) as systematic patterns are implemented throughout the lexicon. In later development, infants begin to shift towards a segment-based approach to word-adaptation, as individual segments rather than whole words are adapted to facilitate the production of adult word forms (Vihman (1996)). This highlights an important step in the course of phonological development, as this allows lexical development on a phoneme-by-phoneme basis, rather than word-by-word (Smith (2010)).

Templates such as consonant harmony have been found to occur universally (Vihman (1978)), yet individual infants have been found to differ substantially in their use of templates (Vihman et al. (1994)), due to factors such as biological constraints, ambient language effects and motor and cognitive factors, as well as the variability in each infant's own input. According to Vihman:

> Each child, drawing on his or her *unique genetic inheritance* to develop vocal resources in the course of the first year of life, *influenced or shaped by the speech of the caretakers*, faces and solves a range of word production problems, resulting in an *idiosyncratic* first phonological organization.
>
> (Vihman 1993:62; emphases added)

But what does this mean for children who share a genetic inheritance and learning environment? With the above statement in mind, we could hypothesize that the phonological development of a set of twins would be similar, or at least show more resemblance than two unrelated children growing up in different learning environments. Though there are very few studies observing the phonological development of twins in the current literature, contrasting reports on this hypothesis can be found in studies by Bloch (1921) and Leonard et al (1980). Bloch discusses briefly the language development of his twin nephews, whose language is "identical: as soon as a modification or an acquisition appears in one infant, the other follows almost immediately" (p. 693, own translation). However, Leonard et al (1980) found that there is no similarity between twins in the acquisition of phonemes, despite the shared learning environment and genetic make-up of the infants. This study will consider this hypothesis further, by focussing on both the similarities and differences in a pair of dizygotic aɔɛtwins acquiring French in the same environment, where it is fair to assume that they are receiving an input of the same nature from their caregivers.

2. *Methodology*

The data used in this study is taken from the CHILDES database (Hunkeler (2005)); a longitudinal study of dizygotic twins, Camille and Pierre, between the ages of 1;3 and 2;2, acquiring French as their target language. The data was taken from a set of bi-weekly sessions, in which the output of each child was transcribed during interaction at home with the mother, who prompted speech using stimuli such as toys or picture books. Only one twin was present during each session, ensuring that the infant's output was not influenced by speech from his or her twin. The data was transcribed by the investigator, who was analysing the early lexical interactions between an infant and his or her mother for his doctoral thesis. As these twins are not genetically identical, any similarities in their language development must either reflect fundamental elements of the acquisition of French as a target language, or 'shared idiosyncrasies' resulting from a shared learning environment and input.

The majority of tokens produced by both children were nouns relating to everyday objects such as animals and items of clothing. Due to the nature of the French noun, which rarely appears without an article, often the children's tokens were preceded by an open vowel, either /a/, /ɔ/ or /ɛ/, creating what could be interpreted as different realizations for the same word type (e.g. in Pierre's data, 'papa' *daddy* had eight different realizations, all with the same stem, including /apapa/, /papa/, /ɔpapa/ and /ɛpapa/). This 'pre-utterance vowel' was also observed in Macken's (1979) study of Si acquiring Spanish, and following Macken's interpretation of Si's vocabulary, this study will count different instances of vowel + noun combinations as the same token, as long as the noun itself remains constant.

This data will be used in this study to observe any similarities in the phonological development of the twins. If an infant's first output is "shaped by the speech of the caretaker", as Vihman (1993, p.62) proposed, then it is possible to hypothesize that these twins will indeed share some similarities in their first word forms that go beyond the universal output patterns identified across all languages (Vihman (1978), Smith (1975)). A longitudinal study of each infant's lexical inventory will identify any recurrent templates in the output, which will be identified as patterns which are commonly used to adapt output forms. These will then be compared speculatively to observe whether there are any inherent similarities between the phonological development of these two infants. The infants' development will be compared with regard to both age and output. A comparison of the infants' development with regard to age will give a clear indication of any differences in the level of their output ability, while a comparison of the two outputs at the 15- and 25-word points (the points at which the infant produces 15 or 25 words in one recording session, respectively) will provide a fair comparison of the infants' lexicons over the course of their development.

3. Analysis

Each infant's output was observed longitudinally and recurrent patterns in both selected and adapted word forms were identified. Both infants were found to produce a large proportion of words with consonant harmony, and both omitted syllables in many forms to produce a simple CV structure. These patterns correspond with Smith's (1975) theory of universal tendencies in phonological development, reflecting universally common tendencies of simplifying word forms.

With regard to age, the twins' development shows no similarity; Camille produces many word forms even in her first session, some of which are realized accurately, while Pierre's output progresses more slowly, and retains a relatively low level of accuracy throughout the study. The twins' outputs in the first session at 1;3 can be seen in Table 1, below.

Camille		**Pierre**	
auto ‘car’	[tu] [tɔ]	*balle* ‘ball’	[bab]
ça c’est un poteau ‘there’s a post’	[sa sɛ œ̃ pɔtɔ]	*bateau* ‘boat’	[ba]
c’est ça ‘that’s that’	[sɛ sa]	*maman* ‘mummy’	[mama]
chat ‘cat’	[ʃa]		
coucou ‘darling’	[tutu]		
maman ‘mummy’	[mama]		
ours ‘bear’	[dɔ]		
sauterelle ‘grasshopper’	[isøtø]		
tchou tchou [train noise]	[tɔʃ tɔʃ]		
tortue ‘tortoise’	[tate], [tata], [toty]		

Table 1: Camille and Pierre’s output forms at 1;3

From Table 1 it is clear that Camille and Pierre differ vastly in their approach to language production. An analysis of each child’s individual phonological development will therefore provide a framework from which the results can be observed, in order to identify any similarities between the two twins.

3.1. Camille

In her first recording session at 1;3 Camille produces a small yet relatively accurate set of word forms, as shown in Table 1. Most of these first words are selected, as Camille produces words which already contain her preferred output patterns, exploiting existing patterns such as consonant harmony in *maman* ‘mummy’ – [mama] and *tortue* ‘tortoise’ – [tata]. Some of her output forms do not match the target word, however, and are ‘adapted’ to fit her capabilities; syllable deletion is seen in *auto* ‘car’, which is reduced to [tɔ], and velar fronting occurs in *coucou* ‘darling’ as the velar plosive is realized as an alveolar plosive [tutu]. A notable occurrence in this first set of words is seen when Camille produces a string of words, *ça c’est un poteau* ‘there’s a post’, produced accurately as [sa sɛ œ̃ pɔtɔ]. Disyllables are the most common structure amongst the first words of infants acquiring French, and indeed they make up the majority of Camille’s word forms throughout the recordings in this study. However, though polysyllables are not uncommon in French as they are in English (Vihman (1991)), this five-syllable word string does suggest that Camille is developing a good phonological memory and an ability to plan her output at a very early stage in her development.

Camille continues with a high level of accuracy in her output, and variability in her word forms is rare, with only 8% of her accumulative word forms produced with any variation by the 25-word point at 1;8. One of these variable forms marks an important development in Camille's phonological progress: in her third session, at 1;4, Camille produces an almost accurate realization of the target *attention* 'watch out', as [sayːtɑ̃sjɔ]. However, at 1;6 this form undergoes syllable deletion to become [tɑ̃sjɔ], outlining an instance of regression in the quality of Camille's output. From this point in her development, Camille's word forms become much more systematic, as she develops a set of preferred output patterns and begins adapting words throughout her inventory to match a set of templates. For example, 'palatal patterns' (replacement of certain segments, such as liquids, with the palatal [j], found to be common in the acquisition of French (Wauquier-Gravelines & Suet-Bouret (2004)), become much more prevalent, as 31% of Camille's output forms contain [j] between 1;6 and 2;0. These forms are either selected due to an existing [j] in the target word, or adapted using [j] in place of segments such as /l/ and /ʀ/, as in the examples in Table 2 below:

Replacement of /l/	**Replacement of /{/**
langue 'tongue' [jɑ̃g] (1;8)	*souris* 'mouse' [sɔji] (1;4)
de l'eau 'water' [djo] (1;8)	*au revoir* 'good bye' [aja] (1;5)
lampe 'lamp' [ja] (1;8)	*roue* 'wheel' [ju] (1;9)

Table 2: use of /j/ in Camille's adapted word forms

In addition to this palatal pattern, over one third of word forms contain a consonant harmony template in this period; Camille is certainly implementing systematic templates to her word forms at this point, facilitating her word learning and maximising her output, which is growing at every recording session.

Just before Camille reaches two years of age, a further development can be seen in her output forms. At 1;11 she finally produces *attention* accurately, as [atɑ̃sjɔ̃], and the appearance of palatal patterns diminishes as the target consonant [l] appears more regularly in her inventory, both in the correct context – *il pleut pas* 'it's not raining' [iplœpæ] (2;0), and incorrectly in *citron* 'lemon' [sitlɔ̃] (2;2) – as she starts to over-generalize this newly stable addition in her phoneme inventory, now replacing target /ʀ/ with [l] instead of [j]. At the age of 2;0 it appears that Camille's word forms start to shift towards a more accurate production, as she begins to adapt individual phonological segments rather than whole words. This change to a segmental approach to production occurs as her output becomes less systematic and moves closer to the adult target.

3.2. Pierre

With regard to his output forms and his approach to word learning Pierre's phonological development differs substantially from that of his twin sister; he produces only four words in his first recording session compared to eleven words in Camille's first session (shown in Table 1). While Camille expands her output very gradually, increasing her production and reaching the 15-word point at 1;7, Pierre produces an average of only four words per session for the first six sessions, and then suddenly increases his output three-fold, reaching his 15-word point a month earlier than Camille, at 1;6. His learning continues in this fashion, increasing his output suddenly and substantially, until he reaches his 25-word point, almost 3 months later than Camille, at 1;11. He displays a large amount of variability in specific word forms, especially in sessions where his vocabulary has shown a large increase since the previous session, as in the following examples in Table 3, taken from the session where he reaches his 15-word point:

***balle* 'ball'**	***chausson* 'slipper'**
[ba]	[ʃøʃɔ̃]
[baj]	[ʃəla]
[bɔ]	[ʃø]
[bajaj]	

Table 3: variation in the realization of the target words *balle* and *chausson* in one recording session.

This variability could reflect instability in Pierre's word form representations, showing that he could be struggling with motor and planning capabilities in his production.

Disregarding the use of the pre-utterance filler, Pierre's word forms are all mono- and disyllabic throughout the recordings, with one exception [apapabrumbrum] – *papa brum brum* 'daddy brum brum' in his final session at 2;2. Alongside a preference for consonant harmony before the 25-word point, which appears in 25% of word forms, both selected and adapted, a small number of output constraints are also applied to Pierre's output forms. He selects words with a shared place of articulation, as in [pɔm] *pomme* 'apple' (1;7.27) and [dɔn] *donne* 'to give' (1;8), or a front-back melody, as in [bato] *bâteau* 'boat' (1;10) and [abɛt] *bêtes* 'creatures' (1;11), both which contain a bilabial-alveolar melody. When target forms do not fit either of these preferred patterns, he adapts them using typical processes such as consonant harmony and syllable deletion, shown in the examples below:

(1) [kaka] *canard* 'duck' – consonant harmony (1;6)

(2) [bibɔ̃] *biberon* 'bib' – syllable deletion (1;11).

Though Pierre's output could be interpreted as less advanced than that of his twin sister in terms of both quantity and quality, his capabilities in terms of his articulatory output are not as restricted as Camille's. Whereas Camille systematically implements a palatal template to deal with the difficult phonemes /l/ and /ʀ/, Pierre is able to produce both of these in some instances, seen in Table 4:

/l/	/ʀ/
[vwala] *voilà* 'there we go' (1;8) [adlo] *de l'eau* 'some water' (1;7)	[ʀɛzɛ̃] *raisin* 'grape' (2;0) [ʀɔ̃] *rond* 'round' (2;1)

Table 4: instances of /l/ and /ʀ/ in Pierre's word inventory

Pierre's use of these difficult segments is by no means across-the-board, and he does use some processes to adapt words which are too challenging, for example:

(3) [baj] *balle,* 'ball' (1;6)

(4) [labɑ] *là-bas,* 'over there' (1;8)

(5) [pɑja] *voilà,* 'there' (1;10)

The lack of a consistent palatal template like Camille's for dealing with these segments reflects an important difference in the twins' development, showing that, though the twins share many of the same phonological challenges in their acquisition of French, they do not share the same methods for dealing with these. However, the example of *voilà* at 1;10, above, shows an occurrence of regression in Pierre's word forms when compared with the accurate realization of *voilà* in Table 4, [vwala], at 1;8. The use of both cluster reduction and palatal gliding in [pɑja] two months later shows evidence towards the use of a systematic output pattern in Pierre's vocabulary, though this may not occur across-the-board. This indicates that Pierre could also be moving towards a segmental approach to word learning, coinciding with Camille to reach a turning point at 2 years.

4. Identifying Templates in the Infants' Outputs

Both twins show signs of reaching a 'turning point' in their phonological development at around 2 years. Accuracy in the output increases as the twins begin to produce challenging segments correctly, and previously incorrect or regressive forms move towards a more accurate realization of the target word. This turning point is best observed through the example of templates, as they develop, regress and move towards the target production over time. The templates identified in these infants' output do not reflect a restriction in the production of individual consonants; rather they show restrictions in the production of word forms. Both twins are able to produce a large variety of phonemes, and no specific phonemes are completely omitted from either twin's inventory. This highlights the existence of a whole-word approach to production, not on a phonetic level, but on a phonetic level.

4.1. Consonant Harmony Template

The twins each implement two templates which are striking in their resemblance. The first template to appear in both infants' output forms is based on consonant harmony, which is prevalent in both selected and adapted word forms for both twins. The production of the fricative /ʃ/ is a challenge for both Camille and Pierre, and the twins use the same consonant harmony adaptations to overcome the difficulties encountered in production of this phoneme.

Though Camille does not consistently produce /ʃ/ as [s], the use of this phoneme certainly limited in her output forms, as can be seen in Table 5.

Correct Realizations of /ʃ/	**Incorrect Realizations of /ʃ/**
du chocolat 'some chocolate' [dyʃojaja] (1;10)	*chapeau* 'hat' [po] (1;9)
cacher 'to hide' [taʃe] (1;7)	[azø, apo] (1;10)
chercher 'to look for' [ʃeʃe] (1;8)	*château* 'castle' [to] (1;8)

Table 5: realization of target /ʃ/ in Camille's inventory

Table 5 shows Camille's attempts at producing this phoneme as well as word forms which are adapted to avoid it, and it is clear that it is in some way limited in her output. Upon closer observation of the distribution of /ʃ/ in Camille's output forms, it is found that she does produce this consonant regularly, but only in instances of consonant harmony. There are only two exceptions to this rule, *du chocolat* and *cacher*, seen in Table 5. More examples of this restriction can be seen in the following word forms, which are very similar in structure, and Camille is able to take advantage of this by implementing a consonant harmony template which enables her to adapt all three targets for production:

(6) a. *chaussures* 'shoes' - [ʃoʃy] (1;7 – 1;9)
b. *chaussons* 'slippers' - [ʃoʃɔ̃] (1;7)
c. *chaussettes* 'socks' - [ʃoʃɔ̃] [ʃoʃɛt] [ʃoʃɛʃ] (1;6 – 1 ;9)

There is an interesting use of consonant harmony at work in this case, as the consonant on the unstressed syllable is harmonized, rather than the /s/ on the stressed syllable. However, at 2;0, *chaussures*, which was produced consistently as [ʃoʃy] for three months, is realized as [sosy]. The shift from harmony of the unstressed syllable to harmony of the stressed syllable indicates that Camille is paying attention to the segmental building blocks of words, even though the output form remains incorrect. Little by little, this phoneme starts to appear correctly in some output forms, as she acquires new forms with /ʃ/ such as [aʃjɛ̃] *chien*, 'dog' (2;0) and [ʃənij] *chenille*, 'caterpillar' (2;2), as well as updating previously inaccurate forms with this new phoneme: [to] (1;8) → [ʃato] (2;1) *château*, 'castle'.

Pierre also appears to struggle with the phoneme /ʃ/ in his output, as well as its alveolar counterpart, /s/. Before 2;0 there are only three realizations of /s/ in his recordings which do not appear in the context of consonant harmony, while /ʃ/ is realized only in situations of consonant harmony. Though Pierre's template is not identical to Camille's, he adapts the same word forms with the same process, to overcome the same phonological restrictions:

(7) a. *chaussures* 'shoes' - [aføfø] (1;8)
b. *chaussons* 'slipper' - [ʃøʃø] (1;4 – 1;10)
c. *chaussettes* 'socks' - [sosɛt] (1;9 – 2;1)

Like his twin sister, Pierre begins to adapt the template as he moves towards a segmental understanding of word forms; [ʃøʃø] becomes [sosɔ̃] at 1;10, and [aføfø] is realized as [ʃoʃɔ̃] at 2;1. Though the output forms are not the same for both twins in all cases, these examples suggest that one twin could be copying the other's output. Due to the limited appearance of /ʃ/ and the subsequent adaptations of words containing this phoneme, it is evident that both infants are struggling with the same output limitations. As one twin begins implementing a consonant harmony template in order to deal with this phoneme, it could be that the other twin takes advantage of this to overcome his or her own difficulties. The template first appears in Pierre's word inventory at 1;4, and two months later it is recorded in Camille's output forms, becoming a rather influential template over her whole lexicon in the following months, as it causes forms to regress through overgeneralization: *abeille* 'bee', first realized accurately as [abej] at 1;7, is realized as [ʃabeʃ] at 1;8.

Like Camille, Pierre's representation of /ʃ/ stabilizes at around 2;0, and both /ʃ/ and /s/ are produced accurately more often:

(8) a. *chaud* 'hot' [ʃo] (2;0)
b. *chaise* 'chair' [ʃɛz] (2;1)
c. *soleil* 'sun' [sɔjɛj] (2;1)
d. *sous* 'under' [su] (2;1)

Once again, this suggests that the twins are paying attention to individual segments as their phonological inventories grow, and they are both reaching this important landmark in their phonological development at around the same time, though it is clear that their development is not intrinsically similar in any significant way.

4.2. /m/-Replacement Template

Another limitation which appears in the phonology of both twins involves the phoneme /m/, and a template is used to adapt some target forms to facilitate production, hereon referred to as the '/m/-replacement template'. Though /m/ is produced accurately from the first recording session (in [mama] *maman* 'mummy' for both twins), this phoneme does not appear often in either twin's initial output, and Pierre's realizations of /m/ are especially limited, only appearing correctly in word-initial position in four target words before 2;0:

Camille's Realization of /m/	**Pierre's Realization of /m/**
maman 'mummy' [mama] (1;3)	*maman* 'mummy' [mama] (1;3)

main	'hand'	[mɛ̃]	(1;7)	*miaou*	cat noise	[mjau]	(1 ;8)
miaou (cat noise)		[mja]	(1;7)	*main*	'hand'	[mɛ̃]	(1;10)
Mini		[mini]	(1;7)	*moi*	'me'	[ma]	(1;11)

Table 6: realizations of /m/ in each twin's lexicon

Pierre continues to implement the /m/-replacement template until the end of the recordings at 2;2. Camille's realization of /m/ is also limited, as she only produces four target words with /m/ before 1;10. However, after this point her use of target /m/ increases rapidly, and she builds on her existing lexicon with words containing this new stable phoneme. As seen in Table 6, both infants are able to produce /m/-initial words when the second consonant in the word is either a nasal or a glide; otherwise, /m/ is replaced with a bilabial consonant with oral stricture. This suggests that both infants have trouble changing from nasal to oral stricture across a word, showing evidence towards a whole-word approach to word learning. The /m/-replacement template adapts words with changing stricture in the target form to facilitate word production, and is found in both twins' output forms. As with the consonant harmony template, the twins' adapted forms are not all identical, but both the phonological challenge and the process used to overcome this are exactly the same. This template first appears in Camille's lexicon at 1;4.21, when she realizes *mouton* 'sheep' as [bɔtɔ]:

Camille				**Pierre**			
mouton	'sheep'	[bɔtɔ]	(1;4)	*morceau*	'piece'	[bøzo]	(2;1)
				monsieur	'Mr.'	[bəsjø]	(2;2)
maison	'house'	[abɛzɔ̃]	(1;11)	*maison*	'house'	[abɛzɔ̃]	(2;0)
manteau	'coat'	[pɔto]	(2;0)	*manteau*	'coat'	[əpoto]	(2;2)
merci	'thank you'	[besi]	(1;7)	*merci*	'thank you'	[besi]	(2;1)

Table 7: use of the /m/-replacement template in each twin's lexicon

Table 7 shows all instances of the /m/-replacement template in the recordings. Pierre's use of this template is more consistent, as it appears more densely in his word inventory; more forms are produced in a shorter space of time. However, the template first appears in Camille's output forms, and as can be seen from the table, identical forms are produced by both twins, Pierre's appearing no more than two recording sessions after Camille first uses the forms. This could be another example of one twin 'borrowing' templates from the other; once again, the twins are meeting the same phonological challenges, and using the same methods to overcome these challenges.

The use of this template provides another good example of the shift from whole-word learning to segmental learning at the two-year stage. This is demonstrated in Table 8, below, where Camille's realization of *maison* 'house' is shown longitudinally:

1;9	1;10	1;11	2;1	2;2
[bɛzɔ̃] [mɛzɔ̃]	[amɛzɔ̃]	[bazɔ̃] [abɛzɔ̃]	[mɛso]	[mɛzɔ̃]

Table 8: development and regression in Camille's realizations of *maison* 'house'

As can be seen in Table 8, Camille is able to produce *maison* accurately at 1;9, with /m/ in word-initial position. However, the representation is still unstable, and in the same month her existing /m/-replacement template is used to adapt this form. In the following month it seems that *maison* has resisted adaptation, but by 1;11, two different realizations of *maison* using this template prove otherwise. The stressed second syllable of the target word continues to be produced correctly, however, as the first syllable is adapted to retain oral stricture throughout the word; Camille is therefore working on words as whole units at this stage. By 2;1, she has started to move from the whole-word stage to learning individual segments of words, and this is reflected in her production of *maison*; the word is no longer adapted with the template, and the word-initial /m/ is realized correctly, showing that Camille is able to shift from nasal to oral stricture within a word. However, the stressed syllable is no longer accurate: the fricative is unvoiced and the vowel is not nasalized, and so as one element of her phonology develops, another regresses. However, this regression does not last long, as by 2;2 Camille is able to realize *maison* as the target form, and by this stage 71% of Camille's /m/-initial targets are produced correctly.

4.3. From Whole Words to Segments

Systematization of the lexicon through the implementation of templates provides a way for the developing infant to practice using words, despite lacking some of the necessary tools to do so. As Thelen and Smith (1994) explained in their theory of cognitive development, dynamic systems theory, action is a necessary part of development, and in the case of linguistic development, 'action' is manifested in word production. Infants select templates based on their own abilities and phonological preferences, and adapt whole word forms to meet their output constraints, as seen in this study, as the production of fricative consonants or oral stricture across whole word forms constrain the output of both Camille and Pierre. As an infant's phonemic lexicon grows, his constraints upon the whole word diminish, and he begins attending to the individual segments which make up the word. As seen in Table 8, this often causes previously accurate segments to regress as attention shifts to different units in the word, but as this example also shows, a regression in accuracy signals development, and accuracy will follow as a result of this.

At 2 years, both Camille and Pierre appear to reach a turning point in their development which marks the shift from whole word learning to segmental learning. Their routes to language in the preceding months show no similarities in terms of ability; the infants approach language learning in different ways throughout the recordings, and their strengths and weaknesses differed as a result of this; Camille's lexicon grew steadily over the recording sessions, while Pierre's remained constant before the occurrence of an apparent vocabulary ‘spurt’, for example. While Camille approached word use carefully, paying attention to detail and accuracy without producing any variability in her forms, Pierre’s approach was comparatively haphazard, with many variable word forms and no real systematic approach to the more difficult phonemes. With this in mind, their joint arrival at segmental learning at around 2 years is surprising, though this is by no means an absolute in their production, as development is still very much on-going at this point. However, this raises questions about the relationship between the language development of a pair of twins: at first so different in nature, how can two infants reach segmental learning at the same time, gathering along the way the same set of templates and even some identical output forms? Further investigation into the development of twins is needed if this question is to be answered, both to identify any similarities in output forms, and to pinpoint any parallels in the landmark developments that infants make on their route into language.

5. Discussion

The findings from this study support the contrasting observations of both Bloch (1921) and Leonard et al (1980). In line with Leonard et al’s findings, the twins do not acquire the same phonological units in the same order; in fact they do not even show signs of using the same approaches to phonological development. However, Bloch’s observation of one twin following the other’s phonological acquisitions is found in the data, as the twins implement the same template which causes them to produce some identical output forms, and no doubt a larger set of recordings would back up this observation even further.

It is unsurprising that the twins are encountering the same phonological difficulties in their outputs. Indeed, the four universal tendencies of phonological systematization (Smith (1975)) arise from a selection of common output constraints encountered during language development, which result in the same universal output patterns. Thus the appearance of consonant harmony in order to deal with a change in manner of articulation, or even the use of oral harmony to avoid changing vocal stricture within a word cannot be considered as unusual, for twins as well as for unrelated infants. However, questions remain regarding the identical forms found in both twins’ outputs; are these identical due to both twins implementing their shared templates on the same words, or is one twin imitating the other’s output forms? As different output forms for the same target words occur in each infant’s lexicon as a result of these templates, it can be deduced that both templates are active in each infant’s phonological development. However, the influence that the twins have on each other’s input and output with regard to these templates cannot be judged from this data, and more recordings would be needed to track the full extent of these findings. It seems that a form of phonological ‘bootstrapping’ is at work here, as one twin uses the other twin’s phonology to overcome his or her parallel phonological challenges. Even with a complete set of data it would not necessarily be possible to answer the question of which came first: the template or the output form – did the infant borrow a template, producing the same output forms as a result of this, or did the infant borrow an output form, which went on to become a template as a phonological hurdle became easier to overcome?

As for similarities in the phonological development of these twins, while these identical templates do not suggest any genetic influence on the twins’ language developments, they do highlight the influence of the learning environment and the interactions within that environment on an infant’s unfolding phonological system. Furthermore, the underlying phonological challenges reflect similarities in the twins’ overall production capabilities, which is backed up further by evidence of a coordinated shift from whole word learning to segmental learning despite a vast difference in the initial approach to language acquisition.

All in all, these findings raise some important questions regarding the language development of twins, especially those concerning the nature versus nurture dichotomy, which remains a contended issue in the world of language acquisition. Similarities found in this study could be down to inherent similarities in the infants' development due to genetic inheritance, though the findings lean towards the influence from a shared learning environment. After all, the learning environment of a pair of twins differs significantly from that of non-twin infants, both in terms of attention from the caregiver as well as input, which will consist largely of the other twin's output forms. All of these criteria must be considered when observing any form of twins' development, and without a large sample of both twin and non-twin infants, hypotheses regarding similarities are unlikely to be proven. However, the similarities found in this longitudinal observation of twins' output forms provides a basis for further investigation, as the implementation of templates sets up an infant for his continuing phonological development, and so the use of identical templates in the output could have some important effects on further language production.

6. References

BLOCH, O. (1921). Les premiers stades du langage de l'enfant [The first stages in an Infant's Language]. *Journal de psychologie [Journal of Psychology].*

EDWARDS, J., BECKMAN, M. E., AND MUNSON, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*, 47, 421–436.

HUNKELER, H. (2005). Aspects of the evolution of the early lexicon in the interactions mother-child: Case study of two dizygotic twin children between 15 and 26 months. University of Rouen.

LEONARD, L. B., NEWHOFF, M. & MESALAM, L. (1980) Individual differences in early child phonology. *Applied Psycholinguistics*, 1. 7-30.

MACKEN, M. (1979) Developmental Reorganization of Phonology: A Hierarchy of Basic Units of Acquisition. *Lingua,* 49. 11-49.

SMITH, C. E. (2010). Exploring the Role of Systematization in Phonological Development: A Dynamic Systems Perspective. MA Dissertation: University of York.

SMITH, N. V. (1973). *The Acquisition of Phonology: A Case Study.* Cambridge: Cambridge University Press.

SMITH, N. V. (1975) Universal Tendencies in the Child's Acquisition of Phonology. In B. Lust and C. Foley (eds.), *First Language Acquisition: The Essential Readings.* London: Blackwell.

THELEN, E. AND SMITH, L. B. (1994). A Dynamic Systems Approach to the Development of Cognition and Action. Cambridge: MIT Press.

VIHMAN, M. M. (1978). Consonant Harmony: Its Scope and Function in Child Language. In J. Greenberg, C.A. Ferguson & E.A. Moravcsik (eds.), *Universals of Human Language, 2*. Stanford, CA: Stanford University Press.

VIHMAN, M. M. (1981). Phonology and the development of the lexicon. *Journal of Child Language*, 8, 239–264.

VIHMAN, M. M. (1991). Ontogeny of Phonetic Gestures: Speech Production. In I.G. Mattingley and M. Studdert-Kennedy (eds.) *Modularity and the Motor Theory of Speech Perception*. New Jersey: Lawrence Erlbaum Associates Inc.

VIHMAN, M. M. (1993) Variable paths to early word production. *Journal of Phonetics,* 21. 61-82.

VIHMAN, M. M., KAY, E., DE BOYSSON-BARDIES, B., DURRAND, C. AND SUNDBERG, U. (1994). External Sources of Individual Differences? A Cross-Linguistic Analysis of the Phonetics of Mother's Speech to 1-Year-Old Children. *Developmental Psychology*, 30, 5, 651-662.

VIHMAN, M. M. AND CROFT, W. (2007). Phonological development: toward a 'radical' templatic phonology. *Linguistics*, 44-5, 683-725.

VIHMAN, M. M. AND VIHMAN, V-A. (in press). From first words to segments: A case study in phonological development. In E. V. Clark and I. Arnon (eds.) *Title TBA, Trends in Language Acquisition Research*. John Benjamins.

WAUQUIER-GRAVELINES, S. AND SUET-BOURET, D. (2004). Acquisition des attaques syllabiques et 'palatal patterns'. Pourquoi yod ? [Aquisition of Syllabic Onsets and 'Palatal Patterns'. Why yod?] *'Phonologie' conférence ['Phonology' Conference]*. Orléans, France. 3rd June 2004. Université de Nantes.

Acknowledgements

With thanks to Marilyn Vihman, Daniel Laing and Edward Haynes.

Catherine Smith
email: ces241@hotmail.com

THE ACQUISITION OF CONSONANT CLUSTERS IN POLISH: A CASE STUDY

MARTA SZREDER

Abstract

This paper examines the phonological processes affecting consonant clusters in the speech of a child acquiring Polish (1;5-1;9). Word-initial, word-medial and word-final clusters are discussed, and compared to word-initial singleton consonants in the data. The nature of the processes, as well as the wide range of variability within the child's system, lead to the conclusion that articulation, attention and word-based processing are the main factors affecting the child's production.

1. Introduction

There is an ongoing debate in the literature on whether phonological processes used by children reflect a re-organising rule-based system, or a system that is emerging from the interaction of developing motor and cognitive skills. An example of the latter approach is presented in Studdert-Kennedy & Goodell (1993), who found that their data pointed to articulatory factors being the main source of child errors. According to Studdert-Kennedy & Goodell, "a child's errors in early words can arise from paradigmatic confusions among similar gestures in a child's repertoire and from syntagmatic difficulties in co-ordinating the gestures that form a particular word" (p. 82). The authors based their conclusions in the Articulatory Phonology framework (Browman & Goldstein 1989; Browman & Goldstein 1992), according to which real-time articulatory gestures are phonological units, stored and produced in meaningful combinations (i.e. words). Under this view, the process of phonological acquisition would consist of learning to produce particular gestures and learning to co-ordinate them into word-shapes. Predicted errors would therefore include paradigmatic and syntagmatic difficulties, just as the ones found by Studdert-Kennedy & Goodell. Such an approach is compatible with the more general whole-word approach in child phonology, first explicitly proposed by Ferguson & Farwell (1975), who claimed that the word is the first basic unit of phonological organisation in children. Word-based processes were demonstrated for other children by several researchers (e.g. Waterson, 1971; Priestly, 1977; Vihman, 1987), who found that children often employ word templates (Vihman & Velleman, 2000) , i.e. patterns of production that change the overall shape of the word rather than affecting particular segments.

A competing approach, rooted in the generative tradition (Chomsky & Halle, 1968), postulates a rule-based (Smith, 1973) or, more recently, a constraint-based (Fikkert & Levelt, 2008; Łukaszewicz, 2007) system as the starting point of phonological acquisition, and assume the development proceeds through re-organisation or re-ranking of the constraints. Under this approach, the basic units of phonological organisation are segments. An important prediction of such an assumption is that all processes should apply to the same segments and the same syllable positions regardless of the word they appear in, i.e. they should be triggered by the same segment-sized unit every time they occur. This is not the case if we adopt the whole-word approach, as the same segments can behave differently in different words. Also, in the constraint-based approach, the processes should not target units larger than the segment, i.e. they should be blind to other positions in the word. Again, the whole-word approach predicts that phonological processes can change the shape of the whole word, as in the case of word templates. Finally, we would like to see the constraints apply at every instance of the particular word being produced, while under the articulatory approach, it is to be expected that attempts at articulating a problematic segment may have different outcomes on different occasions.

To test the above predictions, the current paper examines the behaviour of consonant clusters in the speech of a monolingual child acquiring Polish, with a focus on word-medial clusters in particular. Previous investigations (Zydorowicz, 2007; Łukaszewicz, 2007) have shown that word-medial clusters in Polish are acquired early in development, as compared to word-initial clusters. This has been attributed to various constraint-related factors, from syllable structure constraints (Łukaszewicz, 2007), to morphonotactics and markedness effects (Zydorowicz, 2007.) However, a similar effect has been reported for languages which make use of word-medial geminates (Finnish: Savinainen-Makkonen, 2007) and long consonants (Welsh: Vihman & Croft, 2007; Vihman, Nakai & DePaolis, 2006). The authors of these findings suggest that the relative ease with which the segments are acquired can be attributed to the salience of the word-medial, intervocalic position. Such explanation is further supported by the fact that, in both Finnish and Welsh, word-medial geminates and long consonants often affect word onset in children's production, causing them either to lose accuracy or to be dropped altogether. That consonant clusters can affect the accuracy of other sounds in the word in child language in the same way that geminates and long consonants do has been proposed for Hindi, which makes use of both (Vihman & Croft, 2007). If true for Polish, this might contribute to the explanation of why word-medial clusters are acquired before word-initial clusters. It would also suggest that attention factors play an important role in the acquisition of these segments, and in this sense point to formal constraints not being sufficient to explain the course of phonological development in children.

Taking the above facts into account, the goal of the current paper is to examine the processes affecting consonant clusters in the speech of a Polish child, as well as their effects on other positions in the word. We shall primarily be concerned with three questions (1) how the processes relate to those observable for single consonants, i.e. whether they obey the same constraints; (2) how systematic the processes are, i.e. whether they apply to the same clusters regardless of the word-form and the token they appear in; and (3) whether word-medial clusters trigger instability of word onset, as in Finnish, Welsh and Hindi. It is hoped that the answers to these questions will provide evidence regarding the units of early phonological organisation (words vs. segments), the source of child errors (articulatory vs. formal constraints) and the role of attention in the process.

2. *Method*

The data for the study has been collected from the author's son Grzenio ([gʒɛɲɔ]), a monolingual child acquiring Polish. For the purpose of the current analysis, six half-hour recordings of spontaneous speech in the home environment were selected, with intervals of 12 to 26 days. At the beginning of the study, Grzenio was 1;5.28, with an estimated cumulative vocabulary of about 50 words (MLU 1.2). The recordings ended when he was 1;9.28 and his vocabulary was estimated for about 250 words (MLU 2.6). The total number of interpretable tokens recorded was 1402, and the total number of word forms was 181. As there was no evidence for a qualitative change in Grzenio's phonological organisation, the data will be treated synchronically, with no attention to word form changes over the four months.

To give an idea about the target system, Table 1 presents the consonant inventory of Polish. Consonants marked in grey are those that Grzenio produced at the time of the study. Table 2 (based on data from Milewski, 2005) presents the number of consonant cluster types in different positions in the word in data from preschool children (aged 3-7), spoken Polish, scientific texts and artistic prose.

Table 1: The consonant inventory of Polish.

*the consonant was recorded only once in Grzenio's speech

place/manner of articulation		bilabial	labio-dental	dental	alveolar	palato-alveolar	palatal	velar	glottal	
plosive		p		t				k		-v
		b		d				g		+v
fricative			f*	s	ʃ	ɕ		x		-v
			v	z	ʒ	ʑ				+v
affricate				ts	tʃ	tɕ				-v
				dz	dʒ	dʑ				+v
nasal		m		n			ɲ			+v
liquid	lateral				l					
	rhotic				r					
glide		w					j			

Table 2: Consonant cluster types in Polish (Table based on Milewski, 2005, p. 23)

Position in the word		Preschool children*	Spoken adult Polish**	Scientific Texts***	Artistic Prose***	Grzenio
initial	**N**	249	208	283	310	10
	%	27.85	23.94	22.4	23.0	22.22
medial	**N**	581	588	883	926	27
	%	64.99	67.66	69.9	68.8	60.0
final	**N**	64	73	98	110	8
	%	7.16	8.40	7.8	8.2	17.78

* Milewski, 2005, ** Dunaj, 1985, ***Dobrogowska, 1991

3. *Results*

3.1. Word-initial consonant clusters

As can be seen from Table 2, much fewer clusters occur word-initially than word-medially in Polish. This tendency was also observed for Grzenio, for whom word-initial clusters made up for 22% of all clusters he produced, in line with the frequency observed for preschoolers and adult speakers of Polish (Milewski, 2005). There were only ten types of word-initial clusters in Grzenio's repertoire, all of which are presented in Table 3 along with the targets. It can be observed that all targets are of the structure C1[obstruent] + C2[sonorant], and this structure is preserved in the child form.

Target	Grzenio
br	bw
	bβ
kfj	kx
	tɕl
kl	hj
	kj
	kl
	tɕj
kr	kj
	kl
	kŋ
	tɕj

Table 3: Word onset clusters produced by Grzenio.

However, out of 309 attempts at a target word with a cluster at onset (45 word forms), only in 25 tokens the cluster was not reduced. Moreover, out of those 25, only one was reproduced correctly. Among the remaining 284 child forms, the cluster was reduced to or replaced by a single consonant in 277, and was omitted altogether in 7.

Table 4 presents a selection of target words with consonant clusters in word-initial position along with the child forms. Łukaszewicz (2005) also reports numerous cases of onset cluster reduction, which she finds to be due to either sonority-based deletion, whereby only the less sonorous consonant is retained, or coalescence, where the two consonants are replaced with one, which shares its phonetic properties with both the original sounds. However, Grzenio's forms do not exhibit such consistency, and the process of reduction does not seem to be applied in a systematic fashion. For example, clusters with [s] or [ɕ] as C1 behaved differently in different words. In 'śpi' /ɕpi/ only the stop was retained, in /swɔɲ/ the cluster was replaced with a harmonised consonant, while in 'smok' /smɔk/ it was omitted altogether. The less sonorous stop was retained in /ɕpi/, but it was deleted in 'krab' /krap/, which appeared with an initial nasal palatal [ɲ]. Moreover, different target clusters were often replaced with the same sound, and the initial nasal palatal was also used in the word 'pszczółka' /pʃtʃuwka/. In fact, using coronal and dorsal segments in the place of word-initial clusters was the only pattern that was to some extent regular, in that 50% of the clusters that underwent reduction were either reduced to or replaced by coronal consonants, a further 44% with dorsal consonants, and only 6% with labials (although they were present in 30% of target clusters.)

Table 4: Selected child forms with a word-initial cluster.

*the only form with an accurate word-initial cluster

CC not reduced			**CC reduced/ omitted**		
Target/ Eng	**Target IPA**	**Grzenio IPA**	**Target/ Eng**	**Target IPA**	**Grzenio IPA**
klocki ‘blocks’	klɔtski	tɕjaçka, klɑçki*	klocki ‘blocks’	klɔtski	koçki, tɕaçki
kredka ‘crayon’	krɛtka	kjaxka	Grzenio	gʒɛɲɔ	ɲɛɲɔ, dʑɛɲa,
krab ‘crab’	krap	tɕjapk	krab ‘crab’	krap	ɲap:ka
grzmi ‘thunders’ (V)	gʒmi	bwi	pszczółka ‘bee’	pʃtʃuwka	ɲupkɛ
chrupki ‘crisps’	xrupki	hlupki	słoń ‘elephant’	swɔɲ	nɔɲ, ɲɔɲ
śpi ‘sleeps’	ɕpi	pɕi	smok ‘dragon’	smɔk	ɔŋk
			śpi ‘sleeps’	ɕpi	pi

In sum, although the proportion of word-onset clusters to all clusters in Grzenio’s data was the same as for adult Polish, their production was still very unstable. Obstruent-sonorant combinations were the only ones produced, but most of the time even target clusters conforming to this pattern were reduced. The only word that was pronounced with the correct cluster (‘klocki’ /klɔtski/ > [klaçki]), had as many as five variants, in two of which the cluster underwent reduction. This indicates that even this one instance of correct reproduction of the cluster [kl] was not stable enough to be considered fully acquired. Finally, there was no clear pattern to how given clusters were treated.

3.2. Word-medial consonant clusters

As we have mentioned earlier, word-medial consonant clusters have the largest number of forms among all clusters in Polish (cf. Table 2), and this tendency was also true for Grzenio’s data. Also, while only obstruent-sonorant clusters were used by Grzenio in word onset, word-medial clusters were mostly of the opposite form, i.e. sonorant-obstruent. It must be noted that this tendency is also present in the adult language, where sonorant-obstruent clusters are relatively rare in the word-initial position. In addition, word-medial clusters were produced far more frequently than were the word-initial clusters. Interestingly, the number of word forms Grzenio attempted was similar for both: 45 word-onset cluster word forms (309 tokens) vs. 50 word-medial cluster word forms (270 tokens.) However, the cluster was reduced in only 70 out of the 270 tokens with medial cluster, as compared to 277 for word-onset cluster; also, there was not a single instance of omission. In the remaining 197 child forms the cluster was retained, although it was often reproduced inaccurately. Zydorowicz (2007) reports that morphonotactic clusters seem to be more stable than lexical clusters (i.e. they are rarely modified), but no such tendency was observed in Grzenio’s data, mainly because of a very small number of morphological endings in his speech, presumably due to his young age. In fact, the only morphological suffix in the data that results in a word-medial cluster is the diminutive suffix, and there is no evidence for the productive use of this suffix (i.e. the words only appear in diminutive form.)

The strategy that Grzenio employed to produce word-medial clusters seems to have been more systematic than what we observed for the word-onset ones. Moreover, the substitution pattern was also more strictly defined. All C2s were either coronal or dorsal obstruents, both in the targets and in the child forms. As for C1, it was most often a non-continuant, usually agreeing in place of articulation (PoA) with C2. The exception to both these tendencies was the stop [p], which appeared as C1 in the place of all labial C1 targets, being also the only stop regularly used in this position.

We can therefore extract three main patterns, which are presented in Table 5. All word-medial clusters produced by Grzenio along with the targets are sorted according to C1 (C2 always being a coronal or dorsal obstruent.)

The first pattern applies to all clusters with a labial C1. The preferred C2 is a coronal or dorsal obstruent, and so the cluster /br/ is replaced with /pt/ or /ptɕ/. Still, there is an instance of a cluster without any labial consonant in the target form turned into /pt/.

The second pattern turns all sonorant-obstruent clusters into a sequence of a homorganic nasal and an obstruent. This sequence was also reported as a frequent cluster modification by Łukaszewicz (2007) and Zydorowicz (2007). Again, however, we can see two target clusters that the pattern applies to, despite the fact that they do not match the criteria: /pk/ and /tʃk/, which both become /ŋk/. We can also see that there is a strong preference for coronal and dorsal segments, as the cluster /mp/ is transformed into /nt/.

The third pattern applies to sequences of obstruents, which are transformed into a non-continuant-continuant cluster. This can be seen in the case of the clusters /tk/, /tsk/ and /tʃk/, in which the C2 is reproduced accurately, but C1 appears as several different fricatives.

Approximately half of the child forms which conformed to one of the above three patterns had a cluster of the preferred structure in the target form, and in this sense they were 'selected' (Vihman & Velleman, 2000); the other half of the forms were 'adapted' meaning that the target cluster was transformed to match the pattern. Table 6 presents a selection of child forms with a cluster in word-medial position, sorted according to this distinction.

Again, as was the case with word-onset clusters, we can see that in spite of the general systematicity of the pattern, its application is by no means fully consistent. For example, the word 'Marta' /marta/ is reproduced with two different clusters: [əŋka] and [ɲaɲta], and the same is true for 'soczku' /sɔtʃku/, which appears in two quite different forms: [ɲɔxku] and [øŋku]. Moreover, in the case of the word /spɔdeŋki/, the child form is [dodandi], despite the child's preferred cluster being present in the target form. As regards the 75 word tokens in which the medial cluster was reduced, there was also no obvious pattern as to which segment should be retained: in 32 tokens the cluster was reduced to or replaced with a [-continuant] consonant, but in the remaining 45, with a [+ continuant] consonant.

Interestingly, there were also cases of cluster insertion, where the target form had no cluster but the child form did. Table 7 presents selected child forms with an added cluster.

Table 5: Word-medial clusters produced by Grzenio, sorted by C1.

labial (30 tokens)		**nasal (60 tokens)**		**fricative (98 tokens)**		**other (12 tokens)**	
Target	**Grzenio**	**Target**	**Grzenio**	**Target**	**Grzenio**	**Target**	**Grzenio**
br	pt	jdʑ	ɲdʑ	ɕtɕ	ɕtɕ	tʃk	kk
	ptɕ		ɲt	jɕtɕ	jɕtɕ		tɕk
pk	pk		ɲtɕ	sk	ɕt	xts	jtɕ
ptɕ	pt	lk	ɲk	st	ɕt		
	ptɕ	mp	nt		ɕtɕ		
rt	pt	nd	ɲd		xtɕ		
wk	pk	ɲtɕ	ɲtɕ	ʃk	ɕk		
		ŋk	nd		hk		
			ŋk	ʃtʃ	ɕt		
		pk	ŋk		ɕtɕ		
		rdz	ɲd		htɕ		
		rt	ɲt	tk	çk		
			ŋk		hk		
		tʃk	ŋk		xk		
		wk	ŋg	tsk	çk		
			ŋk		ɕk		
					hk		
					jçk		
				tʃk	hg		
					hk		
					xk		
				tʃn	ɕtɕ		
				xts	çtɕ		
					ɕtɕ		
					htɕ		
					χtɕ		

Table 6: Selected child forms with a consonant cluster in word-medial position.

Select			Adapt		
Target/ Eng	**Target IPA**	**Grzenio IPA**	**Target/ Eng**	**Target IPA**	**Grzenio IPA**
świnka 'pig'	ɕfiŋka	ɕiŋka	spodenki 'trousers'	spɔdɛŋki	dodandi
nie chcę 'not want'	ɲɛxtsɛ	ɲɛɕtɕɛ	zebra 'zebra'	zɛbra	ɲɛptɕa
rybka 'fish'	rɨpka	ɲʔpka	soczku 'juice'	sɔtʃku	ɲɔxku, øŋku
po prostu 'just'	pɔprɔstu	tɔtɔɕtu	pszczółka 'bee'	pʃtʃuwka	ɲupkɛ
nóżkę 'leg'	nuʃkɛ	ɲiɕkə	Marta	marta	əŋka, ɲaɲta
Łukaszka (Gen)	wukaʃka	kaɕka, kahka	kredka 'crayon'	krɛtka	kjaxka
babcia 'grandma'	baptɕa	ɲaptɕa	grzeczny 'good'	gʒɛtʃnɨ	kɔɕtɕi

Table 7: Selected cases of cluster insertion.

Target/ English	Target IPA	Grzenio IPA
buty 'shoes'	butɨ	ʔɲtɕɨ
chce 'want'	xtsɛ	tɕɛŋk
czytać 'to read'	tʃɨtatɕ	tɕiɕtatɕ
dywan 'carpet'	dɨvan	diɲda
idzie 'walk'	idʑɛ	iɕdʑɛ
krab 'crab'	krap	ɲap:ka, tɕjapk
leżeć 'to lie'	lɛʒɛtɕ	jaɕtɕ
Łukasz	wukaʃ	guçkaɕ
oko 'eye'	ɔkɔ	əŋkɔ
tukan 'toucan'	tukan	ɲ:kaɲk

3.3. Word-final consonant clusters

Unlike in the case of word-initial and word-medial clusters, the percentage of clusters Grzenio used in word-final position was relatively higher than observed for adults. Interestingly however, only two types of word-final cluster were used accurately: /ɕtɕ/ and /ɲtɕ/. No other word-final clusters were attempted, and all other clusters present in the data came from cluster insertion, instances of which were presented in Table 8. All the word-final clusters appeared in one of the three forms that were observed for word-medial clusters.

3.4. Word-initial singleton consonants

At the time of the study, the child produced word-initial single consonants with high accuracy, ranging from 75% for labials and 86% for velars to 95% for coronals and palatals. Interestingly, even within this very small margin of variation, there was a difference in the behaviour of the three places of articulation of stops. While the coronals seemed the least variable of the three places, the variability could usually be attributed to articulatory factors, meaning that the sounds often underwent palatalisation, and sometimes even affrication, as in the word ‘tatuś’ /tatuɕ/, which often appeared as [tɕatiɕ].

Labial stops, on the other hand, were rarely affected by segment-based processes (there were nine cases in total of a change in voicing or manner), but, in comparison to the coronals, they were more prone to whole-word processes such as assimilation, resulting in lower accuracy overall. This is illustrated by the word ‘buty’ /butɨ/, usually pronounced [nuta], and ‘babcia’ /baptɕa/, almost always rendered as [ɲaptɕa]. As regards the velar stops [k] and [g], the former was usually pronounced correctly, but the latter was rather infrequent and sometimes replaced with another sound, as in the word ‘gitara’ /gitara/ > [titaja].

In summary, while the stops were rather stable in word-initial position, coronals and palatals were the least susceptible to the influence of other segments in the word, despite being at the same time the least precisely articulated, whereas labials did not undergo many segment-based processes but were often at least partially assimilated to other consonants in the word.

While coronal stops were sometimes replaced with affricates, word-initial affricates often also underwent reduction to stops. Only the palato-alveolar affricates were present in Grzenio’s data and those were usually produced accurately (100% accuracy for [tɕ] and 93% for [dʑ].) Nevertheless, dental and alveolar affricates were palatalised to [tɕ] and [dʑ] approximately half of the time, while at other times being reduced to dental stops (as in the word ‘cześć’ /tʃɛɕtɕ/, pronounced as [teɕ]), but never replaced by any other consonants.

A similar pattern to that of stops was observed for word-onset nasal segments, whose accuracy ranged from 17% for the dental [n] and 66% for the labial [m] to 83% for the palatal [ɲ]. However, in the case of the coronals [n] and [ɲ], the variability was almost entirely limited to the two varying with each other, i.e. [n] was only replaced with [ɲ], while [ɲ] was pronounced as either [n] or sometimes [j]. Again, labial [m] was an exception: almost all of the 34% of inaccurate tokens were instances of consonant harmony (e.g. ‘miś’ /miɕ/ > [ɲiɕ]). That the variation among coronals can be attributed to articulatory difficulties is further confirmed by the behaviour of the glide [j], which not only replaced the nasal [ɲ] in some words, but also was replaced by it in others, although it was produced accurately 86% of the time (for comparison, the labial glide [w] was never used accurately.)

As regards fricatives, they were still relatively undeveloped and infrequent. Apart from a single appearance of [f] during the final session, only palato-alveolar [ɕ] (used interchangeably with palatal [ç]) and velar [x] (used interchangeably with glottal [h]) were produced in word-initial position. Those consonants were also used to replace other fricatives, along with a range of other sounds.

Finally, as is typical for children at his age, except for a single instance of [l] Grzenio did not produce liquids, which he usually replaced with glides.

In summary, on the basis of the behaviour of word-initial segments, we can see that obstruents and glides were the most developed consonants in Grzenio’s data, and among them coronal and palatal segments played a special part. It is perhaps worth noting that the coronals may be particularly difficult in Polish, as the language distinguishes between four different places of articulation for those segments. Not surprisingly, at the time of the study, Grzenio did not yet use all of them. He used dental stops (but not affricates), none of the alveolars, all pre-palatals except for the voiced fricative [ʑ] and both palatals (with the occasional addition of [ç].) Still, the consonants that he produced were often used interchangeably. On the other hand, those relatively unstable sounds, when produced in word onset, were seldom influenced by other positions in the word. In fact, if we compare all inaccurate child forms, articulatory errors (i.e. variable degree of voicing, nasalisation, palatalisation and affrication) make up 80% of tokens with word-initial coronal or palatal obstruents in the target, but only 7.5% and 28% of tokens produced for targets starting with labial and velar obstruents, respectively. The rest of the errors are the result of either omission or assimilation. Table 8. presents instances of omission of word onset.

Comparison of those word forms confirms that labial segments were the most susceptible to variation. Whether it was the glide [w], the nasal [m] or the labial stop [p], they were likely to be omitted. Nevertheless, there were also some cases of omission of coronal and palatal segments, even a consonant as stable as the glide [j]. What seemed to trigger those processes was the presence of a consonant cluster later in the word, as in 'jeszcze' /jɛʃtʃɛ/, reduced to [ɛɕtɕɛ]. In fact, even the variability within the articulatorily motivated range appeared most frequently in words with word-medial clusters. Out of the 23 most variable word types (i.e. the ones for which four or more different child forms were recorded), 12 (52%) had a consonant cluster in word-medial position in the target form. In comparison, out of 39 words which appeared in only one form (but in more than one token), only 7 (17%) had a word-medial cluster.

Table 8: Selected cases of word-initial consonant omission.

Target/Eng	Target IPA	Grzenio IPA
dobranoc 'goodnight'	dɔbranɔts	aɲɛɲɔtɕ
jest 'is'	jɛst	ex
jeszcze 'more'	jɛʃtʃɛ	iɕtɕɛ
łóżko 'bed'	wuʃkɔ	uhkɔ
Łukasz	wukaʃ	ukaɕ
Łukaszek	wukaʃɛk	ukahək
Marta	marta	əŋka
miś 'teddy bear'	miɕ	iɕ
misia 'teddy bear' (Gen)	miɕa	iɕa
piłka 'ball'	piwka	iŋka
pompon 'pompon'	pɔmpɔn	əntən
soczku 'juice'	sɔtʃku	aku

To sum up, while the accuracy of word-initial singleton consonants was very high, the occasional errors that did appear fell into one of two general categories. Firstly, there were errors that could be said to be articulatorily motivated, i.e. resulting from imprecise articulation or, in the case of consonants that had not yet been acquired, from substitution with a similar sound. The processes that fell into this first category all involved variation in voicing, nasalisation, palatalisation and affrication, or from substitution strategies common also in children acquiring English, such as the gliding of liquids (cf. Grunwell, 1985) The second category comprised errors in which the word-initial consonant was replaced with a sound that shared more properties with consonants appearing later in the word than with the target sound. In this sense, those processes seemed word-based. Table 9 presents selected child forms with inaccurate word onset, sorted according to this distinction.

Table 9. Selected child forms with inaccurate word onset, sorted by type.

Segment-based processes			**Whole-word processes**		
Target/Eng	**Target IPA**	**Grzenio IPA**	**Target/Eng**	**Target IPA**	**Grzenio IPA**
bardzo ‘very’	bardzɔ	maɲdʲɔ	babcia ‘grandma’	baptɕa	ɲaptɕa
burza ‘storm’	buʒa	wuɕa	buty ‘shoes’	butɨ	nuta
co ‘what’	tsɔ	tɕɔ	gitara ‘guitar’	gitara	titaja
czapki ‘hats’	tʃapki	tapki	Łukasz	wukaʃ	guçkaɕ
cześć ‘hi’	tʃɛɕtɕ	tɛɕ	Marta	marta	ɲaɲta
czysty ‘clean’	tʃɨstɨ	tɕiɕtɕɛ	miś ‘teddy bear’	miɕ	jiɕ, ɲiɕ
jestem ‘I am’	jɛstɛm	ɲɛɕɛm	po prostu ‘just’	pɔprɔstu	tɔtɔɕtu
koń ‘horse’	kɔɲ	gaɲ	Wanda	vanda	daɲda
leży ‘lies’ (V)	lɛʒɨ	jɛɕi	zebra ‘zebra’	zɛbra	wɛbɛ, ɲɛptɕa
pan ‘mister’	pan	baɲ	zejść ‘go down’	zɛjɕtɕ	jɛjɕtɕ
robić ‘make’	rɔbitɕ	jɔbitɕ, ɲɔpitɕ			
rybka ‘fish’	rɨpka	ɲʔpka			
sam ‘alone’	sam	ɕam			
szafa ‘closet’	ʃafa	hafa			
tatuś ‘daddy’	tatuɕ	tɕatiɕ			
wylał ‘spilled’	vɨlaw	bɨlaw			

4. Discussion

A comparison of the behaviour of consonants in the four positions discussed above (singletons at word onset, word-initial consonant clusters, word-medial and word-final consonant clusters) suggests that each poses different challenges to the child. This is particularly apparent in the case of four consonant types: labials, coronals, dorsals, and fricatives.

We have seen that labial segments were not very stable in the child’s production at the time of the recordings. The fricatives were only emerging, voiced [v] is not present at all, and voiceless [f] was recorded only once, during the last session. The glide [w] did appear, but was never used accurately (only as a substitute for another consonant.) In fact, of all the labial consonants available in the target language, only stops and the nasal [m] were used consistently. However, even these segments exhibited a much higher degree of variability than their coronal and dorsal counterparts, and in particular variability that was not limited to articulatory distortion, but was often the result of assimilation to another consonant. Perhaps not surprisingly, labial stops were also most likely to be omitted in word-onset clusters. We could argue that they still posed difficulties of articulation for Grzenio, and thus were much more vulnerable when co-articulatory factors came into play. Nevertheless, the situation of labial stops was slightly different when they appeared as C1 in word-medial clusters. Specifically, they seemed to be the only stops immune to the cluster template, which replaced all C1 stops with continuants. Highly susceptible to variation at word onset, even if not constituting a part of a cluster, they were almost change-resistant when in syllable coda, even though there was another consonant immediately following that could have been expected to affect them.

The situation of coronal and dorsal segments was very different in this respect. Although the fricatives and liquids were still seldom present at word onset and usually replaced by other segments, obstruents (particularly coronals) and glides, while not always precisely articulated, were very rarely affected by other segments in the word. Moreover, they were also usually retained in word-initial clusters, and in fact the very few clusters that Grzenio produced in this position consisted of a coronal or dorsal obstruent followed by a liquid or a glide. But again, the sounds behaved very differently in word-medial clusters. Here, the obstruents, which were so stable at word onset, were almost invariably transformed into fricatives or nasals whenever they appeared as C1 in medial clusters. On the other hand, C2 in word-medial clusters tended to be coronal or dorsal even if C1 was the labial [p].

As regards manner of articulation, liquids were in general produced only as C2 in word-initial clusters, while fricatives occurred as C1 in word-medial clusters. The behaviour of fricatives here thus confirms the findings of Ferguson (1975), according to which fricatives tend to be acquired first in syllable coda. Moreover, this finding shows that even the constraints on the form of clusters differed depending on word position. First of all, Grzenio mainly produced consonant clusters in word-medial position. This was so despite the fact that clusters are equally frequent in both positions in adult Polish. Also, whereas some clusters preferred by Grzenio are not allowed word-initially in the target language (e.g. [ŋk], [nt]), in other cases the target clusters are structurally the same in both positions but were attempted only word-medially by the child. In effect, the sets of clusters that Grzenio produced in the two positions were mutually exclusive. For example, in the word ‘chce’ /xtsɛ/ ‘want’, the onset cluster was reduced to [tɕɛ]; but when negation was added, so that the cluster appeared intervocalically, it was fully preserved in the child form, the resulting ‘nie chce’ /ɲɛxtsɛ/ ‘not want’ pronounced as [ɲɛɕtɕɛ].

The preferred structure of clusters is also apparent in the templates applied to many of them. Thus, the preferred structure of word-initial clusters is C1[-continuant]+C2[+continuant], while for word-medial position it is C1[+continuant]+C2[-continuant] (with the exception of [p]+[obstruent] clusters.) In fact, there are a few examples of targets which meet both those conditions, and they are produced roughly correctly by the child. For example, the word ‘kredka’ /krɛtka/ is rendered as [kjaxka] and the word ‘klocki’ /klɔtski/ is pronounced as [klaçki].

In general, the constraints on the phonological behaviour of consonants in the data seem to be strongly dependent on their particular position in the word, rather than only on the phonetic identity of particular segments. However, this is not to say that the latter is irrelevant. As we have seen, some of the processes affecting the consonants used by Grzenio appear to have been segment-based, i.e. the variation observed for a given sound could not be explained by an influence of other segments in the word. However, similarly to the errors discussed by Studdert-Kennedy & Goodell (1993), the segment-based processes were always articulatorily motivated, in the sense that the child’s rendition of a given segment was close to the target with respect to its articulatory properties. For example, the initial [b] in ‘burza’ /buʒa/ varied with another labial segment, [w]; [ɲ] varied with [j], which differs only in the degree of closure between the tongue and the palate; and the coronal obstruents appeared with variable degrees of palatalisation and affrication. Moreover, the templates applied to clusters often included consonants agreeing in the place of articulation, which would suggest that articulatory factors might also be partly responsible for its emergence . More specifically, producing a sequence of a continuant and a non-continuant sound with the same place of articulation only in fact requires slowing down the constriction process, and thus is presumably easier to produce for a child than other types of consonant clusters.

Nevertheless, while many processes in Grzenio's data could be explained by imprecise articulation, those processes were more likely to occur under particular conditions. First of all, word-medial consonant clusters showed the same effect as long consonants in Finnish and Welsh, affecting the accuracy of word onset. This suggests that articulatory difficulties were intensified (and sometimes caused) by planning issues and the attentional effort required to produce sounds in a given sequence. Secondly, the fact that the consonant cluster patterns were sometimes applied to consonant clusters irrespective of their structure indicates, that there is more to Grzenio's phonological system than just on-line articulatory difficulties. Specifically, it often seemed that the template targeted clusters on the basis of their abstract property of being a cluster and was not reserved for particularly troublesome combinations of sounds. It was even applied to the same words in different ways on different occasions. While difficulties in articulation can certainly be said to underlie the emergence of the pattern, the strategy employed to deal with those difficulties seems to be based on a generalisation suggesting the presence of an emerging phonological system.

The combination of purely motoric articulatory constraints and articulatorily-motivated yet pre-planned patterns could support the Articulatory Phonology approach (Browman & Goldstein, 1986, Browman & Goldstein 1992), suggesting that the child's phonological organisation is based on articulatory schemata ('gestures' in Browman & Goldstein's terminology.) On the one hand, the approach views gestures as physically real events co-ordinated in real time, which would account for the low degree of systematicity observed in Grzenio's production. On the other hand, they serve as phonological representations, which would account for phonologically determined systematicity on a level higher than that of individual sounds, also observed for Grzenio. Examples (1)-(7) illustrate these two properties of Grzenio's phonology:

1. /buʒa/ > [buɕa], [wuɕa]
2. /miɕ/ > [iɕ], [wɛɕ], [jiɕ], [ɲiɕ]
3. /baptɕa/ > [baptɕa], [japtɕa], [næptə], [ɲaptɕa], [daptɕa]
4. /marta/ > [jaɲta], [əŋka], [ɲapta]
5. /spɔdɛŋki/ > [dɛdɛŋki], [dɔdandi]
6. /klɔtski/ > [tɕjaçki], [klaçki], [tɕaçki], [kɔçki]
7. /krap/ > [ɲapka], [tɕjapk]

In (1) and (2) we can observe the articulatory factors at play, when the initial labials [b] and [m] are substituted by the labial glide [w], although in (1) the accurate form is used as well. However, in examples (2) and (3), there are also forms in which the labial segment assumes a palatal place of articulation, likely due to the influence of the pre-palatal sound later in the word. In addition, the substituted palatal undergoes articulatorily motivated changes as well, when it varies between the glide [j] and the glide [ɲ]. This substituting segment sometimes becomes a coronal, as in (3), and in (2) is deleted altogether. With regard to the word-initial labial, (4) behaves similarily to (3), but at the same time the word-medial cluster appears in three different forms, although each of the forms is compatible with the general cluster template that Grzenio used. The same happens with the word-medial cluster in (5), despite the fact that the target cluster is already of the preferred form, and at the same time the word-initial cluster is reduced and harmonized with the following coronal segment. In (6), the word-initial cluster has the preferred structure, but is sometimes deleted nevertheless. Where it is retained, it appears in two different forms, one of which, [tɕj], is the same as the one in (7), presumably because of the similar target form. However, (7) also has a form with word-initial palatal [ɲ], which is not the case with (6). In other words, each of the transformations can be explained by at least one of the relatively regular processes that were observed for the data set as a whole. Nevertheless, these processes often not only act together, but are also applied in a broadly unsystematic way, making it impossible to postulate any categorical rules for Grzenio's production.

5. Conclusions

Despite the relative systematicity of the phonological processes in Grzenio's data, the nature of the processes as well as the manner of their application seem to suggest, that articulatory and attentional factors are the main source of the child's errors. Moreover, there is evidence that the word, rather than the segment, is the basic unit of phonological organisation for the child. First of all, the processes are triggered by overall shape of the word, meaning that there is notable interaction between initial and medial position, as word-medial clusters affect the stability of word onset consonants. Secondly, the processes affect the overall shape of the word, in the sense that there are templates for articulatory patterns that replace consonant clusters as whole units, rather than follow from the properties of the particular consonants forming the target cluster. Finally, although motivated with regard to articulation, the processes are neither categorical nor obligatory, as they affect only some potential targets and that only part of the time. Therefore, although there are broad regularities in the child's production, the forms are largely unpredictable and resist formulation in terms of any segment-based rules. The observed patterns suggest that the child's phonological organisation is the combination of articulatory and attentional, as well as their interrelations within particular words.

References:

BROWMAN, C. P. & L. GOLDSTEIN 1986. "Towards an articulatory phonology." *Phonology Yearbook*, 3, pp. 219-252.

BROWMAN, C. P. & L. GOLDSTEIN 1992. "Articulatory Phonology: An overview." *Phonetica*, 49, pp. 155-180.

CHOMSKY, N & M. HALLE 1968. *The Sound Pattern of English*. Harper Row, New York.

DOBROGOWSKA, K. 1991. "Grupy spółgłoskowe w różnych stylach języka polskiego." In M. Grochowski (ed.) *Problemy opisu gramatycznego języków słowiańskich*. Warszawa: Instytut Języka Polskiego PAN

DUNAJ, B. 1985. "Grupy spółgłoskowe współczesnej polszczyzny mówionej (w języku mieszkańców Krakowa)." *Zeszyty Naukowe UJ, Prace Językoznawcze* 83

FERGUSON, C.A. & C.B. FARWELL 1975. "Words and sounds in early language acquisition." *Language,* 51, pp. 419-439.

FERGUSON, C.A. 1975. "Fricatives in child language acquisition." In L. Hellman (ed.), *Proceedings of the Eleventh International Congress of Linguists.* Bologna: Mulino.

FIKKERT, P. & C. C. LEVELT 2008. "How does place fall into place? The lexicon and emergent constraints in the developing phonological grammar." In: P. Avery, B. Elan Dresher & K. Rice (Eds.), *Contrast in phonology: Perception and Acquisition*. Berlin: Mouton.

GRUNWELL, P. 1985. *Phonological Assessment of Child Speech (PACS)*. Windsor: NFER-Nelson.

ŁUKASZEWICZ, B. 2007. "Reduction in syllable onsets in the acquisition of Polish: Deletion, coalescence, metathesis, and gemination." *Journal of Child Language* 34(1), pp. 52–82.

MACKEN, M. A. 1979. "Developmental reorganization of phonology: A hierarchy of basic units of acquisition." *Lingua*, 49, pp. 11-49.

MILEWSKI, S. 2005. "Grupy spółgłoskowe w języku mówionym dzieci przedszkolnych. "*LOGOPEDA* 1(1) pp. 5- 32.

PRIESTLY, T.M.S. 1977. "One idiosyncratic strategy in the acquisition of phonology." *Journal of Child Language*, 4, pp. 45-65. Prutting, C.A

SAVINAINEN-MAKKONEN, T. 2007. "Geminate Template: A model for first Finnish words." *First Language,* 27(4), pp. 347-359.

SMITH, N. V. 1973. *The acquisition of phonology: A case study.* Cambridge University Press, Cambridge.

STUDDERT-KENNEDY, M. & E.W. GOODELL 1993. "Acoustic evidence for the development of gestural coordination in the speech of 2-year-olds: A longitudinal study." *Journal of Speech and Hearing Research*, 33, pp. 707-727.

VIHMAN, M. M. & W. CROFT 2007. "Phonological development: Toward a 'radical' templatic phonology." *Linguistics*, 45, pp. 683-725.

VIHMAN, M. M., S. NAKAI & R. A. DEPAOLIS 2006. "Getting the rhythm right: A cross-linguistic study of segmental duration in babbling and first words." In M. C. Pennington (ed.), *Phonology in Context* (pp.25-50). Luton: Macmillan.

VIHMAN, M. M. & S.L. VELLEMAN 2000. "The construction of a first phonology." *Phonetica*, 57, pp. 255-266.

WATERSON, N. 1971. "Child phonology: A prosodic view." *Journal of Linguistics,* 7, pp. 179-211.

ZYDOROWICZ, P. 2007. "Polish morphonotactics in first language acquisition." In F. Menz and M. Rheindorf (eds), *Wiener Linguistische Gazette* 74, pp. 24–44.

Marta Szreder
Department of Language and Linguistic Science
University of York
Heslington
York
YO10 5DD
email: mmsp501@york.ac.uk

HIDEKAZU TANAKA

Abstract

This article shows that fragmentary negative concord answers in Japanese are derived by movement of the negative concord item to the specifier of C with a negative feature. Movement is driven by an uninterpretable feature on the negative concord item. Multiple fragmentary negative concord answers form a surprising constituent (Takano (2002)).

1. Introduction

English allows a fragmentary negative answer that only supplies the value for the wh-phrase in the question in (1). Such a negative answer uses a negative quantifier.

(1) A: Who ate pizza?
    B: Nobody.

While (1) is possible, multiple wh-questions do not allow fragmentary negative quantifiers.

(2) A: Who ate what?
    B: *Nobody, nothing.

Japanese, on the other hand, allows fragmentary negative quantifier answers in both kinds of questions. (3) is an example of a single wh-question, while (4) is a multiple wh-question. They can both be answered by fragmentary NCI answers.

(3) A: Dare-ga pizza-o tabeta-no?
    who-nom pizza-acc ate-Q
    'Who ate pizza?'

    B: Dare-mo.
    Nobody

(4) A: Dare-ga nani-o tabeta-no?
    who-nom what-acc ate-Q
    'Who ate what?'

B: Dare-mo nani-mo.
Nobody nothing
'Nobody, nothing.'

Following Watanabe (2004), we will refer to expressions like *dare-mo* or *nani-mo* as negative concord items (NCIs). This paper shows that fragmentary NCI answers in examples like (3) and (4) are derived by movement and ellipsis, that fragmentary NCI answers observe island constraints, that NCIs in Japanese are adjuncts, and that multiple NCI answers like the one in (4) form a surprising constituent (Takano (2002)). As a theoretical consequence, I argue that ellipsis is conditioned by syntactic identity (Fiengo and May (1994), Merchant (2007), Sag (1976)).

2. Island Effects

Since Chomsky (1977), it has been standard to assume that island conditions (Ross (1967)), or the subjacency condition that puts island conditions together, are conditions on movement operations. (5) is an example of unbounded wh-movement, assumed to be the result of successive-cyclic local movement. This is not possible in (6) due to the wh-island condition.

(5) What do you think that Mary ate?
(6) *What do you wonder when Mary ate?

Japanese also permits long-distance wh-movement, and a fragmentary NCI is also possible in response to such a question.

(7) A: John-ga [Mary-ga nani-o tabeta-to] omotteiru-no?
-nom -nom what-acc ate-comp think-Q
'What does John think Mary ate?'

B: Nani-mo.
Nothing

Nishigauchi (1990) argues that Japanese observes the wh-island condition. The wh-question in (8) is marginal. The fragmentary NCI is even worse than the question itself.

(8) A: ?? John-ga [Mary-ga nani-o tabeta-kadooka] siritai-no?
-nom -nom what-acc ate-whether want-to-know-Q
'What does John want to know whether Mary ate?'

B: * Nani-mo.
nothing

A wh-phrase in a complex NP is perfectly grammatical, but it is impossible to answer such a wh-question with a fragmentary NCI answer.

(9) A: John-ga [nani-o tabeta] hito-o sagasiteiru-no?
-nom what-acc ate-comp person-acc looking-for-Q
'What is John looking for a person who ate?'

B: * Nani-mo.
nothing

The above examples illustrate effects of the wh-island condition and the complex NP constraint on fragmentary NCI answers. They also observe the inner island condition (Ross (1983)). This point is shown by (10).

(10) A: Dare-ga pizza-o tabe-nakatta-no?
who-nom pizza-acc eat-neg-past-Q
'Who didn't eat pizza?'

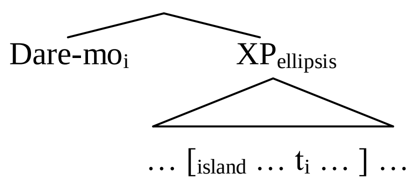
B: * Dare-mo.
Nobody

Thus, not only strong islands, but also weak islands are operative in fragmentary NCIs. The next section proposes an account of this observation.

3. Proposal

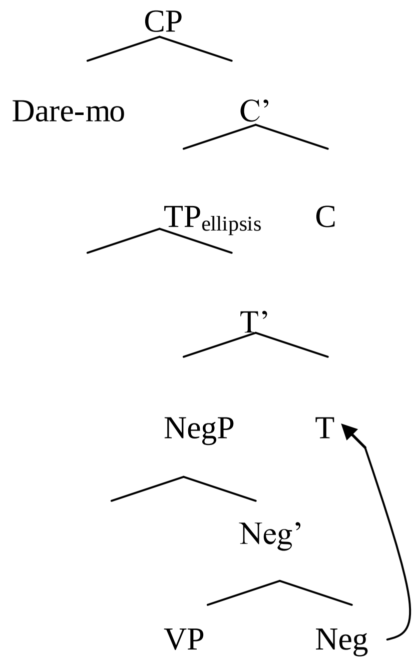
Under the standard assumption that subjacency is a condition on movement, the observation in the previous section suggests that fragmentary NCI answers involve movement. Our conclusion therefore is that fragmentary NCI undergoes movement out of these islands before ellipsis applies, as shown in (11), where $XP_{ellipsis}$ is the elliptical constituent.

(11)



Particularly interesting is the fact that fragmentary NCI answers also observe the inner island condition, as in (10). This shows that the fragmentary NCI in (10) moves out of the scope of the negation. I assume the standard Japanese clause architecture in (12).

(12)



The ungrammaticality of (10) suggests that the fragmentary NCI must move out of the scope of negation. Kishimoto (2005) argues convincingly that the negative head raises up to the T head in Japanese. The fragmentary NCI thus raises out of the TP to a higher projection. The reasonable landing site is the CP-spec. I thus assume that the fragmentary negative quantifier moves to the CP-spec, followed by the ellipsis of the complement of the C head, a TP.

An obvious question is why fragmentary NCIs observe subjacency even after ellipsis. The problem is that ellipsis in English is known to cancel island violations. The grammatical (13) would not be grammatical were it not for the ellipsis of the constituent in angled brackets (Chung et al. (1995), Fox and Lasnik (2003), Merchant (2001), Ross (1969) among others).

(13) I believe the claim that he hit someone, but they don't know who <*I believe the claim that he hit>.

Our observation that fragmentary NCIs observe subjacency therefore seems to be inconsistent with what is observed in English. However, this is not a major problem. It is pointed out by Romero (1998) and Sauerland (1996) that sluicing with an adjunct wh-phrase as a remnant observes subjacency even after ellipsis.

(14) *The governor investigated whether inflation will rebound strongly, but didn't announce how strongly$_i$ <the governor investigated whether inflation will rebound t_i. >

NCIs in Japanese are adjuncts. A piece of evidence for this comes from the following consideration. The answer in (15) shows that the NCI can co-occur with a Case-marked phrase.

(15)	A:	Dare-ga	pizza-o	tabeta-no?	
		who-nom	pizza-acc	ate-Q	
		'Who ate pizza?'			
	B:	Gakusei-ga	dare-mo	pizza-o	tabe-nakatta.
		Student-nom	nobody	pizza-acc	eat-neg-past
		'None of the students ate pizza.'			

The NCI in (15) is similar to a floating quantifier that quantifies the subject.

(16)	Gakusei-ga	san-nin	pizza-o	tabe-nakatta.
	Students-nom	three-class	pizza-acc	eat-neg-past
	'Three students didn't eat pizza.'			

One property of such subject floating quantifiers is that they cannot be preceded by a VP-adverb. Thus, (18) contrasts with (17).

(17)	Gakusei-ga	san-nin	isoide	pizza-o	tabe-nakatta.
	Students-nom	three-class	hurriedly	pizza-acc	eat-neg-past
	'Three students didn't eat pizza hurriedly.'				
(18)	*Gakusei-ga	isoide	san-nin	pizza-o	tabe-nakatta.
	Students-nom	hurriedly	three-class	pizza-acc	eat-neg-past
	'Three students didn't eat pizza hurriedly.'				

I will not try to account for this contrast. What is interesting is that the NCI in (15) patterns with the subject floating quantifier.

(19)	*Gakusei-ga	isoide	dare-mo	pizza-o	tabe-nakatta.
	Student-nom	hurriedly	nobody	pizza-acc	eat-neg-past
	'None of the students ate pizza hurriedly.'				

NCIs thus behave like floating quantifiers, a species of adjunct. If NCIs are adjuncts, the above observation that fragmentary NCI answers observe subjacency is not a problem, since (14) shows that English sluicing also displays island effects when the remnant is an adjunct.

4. Multiple Fragmentary Answers

Let us now focus on multiple fragmentary NCI answers. The account depicted in (12) is based on the assumption that fragmentary NCIs occupy the CP-spec. A question arises is how multiple negative quantifiers, such as the one in (4), are represented. (4) is repeated here.

(20)	A:	Dare-ga	nani-o	tabeta-no?
		who-nom	what-acc	ate-Q
		'Who ate what?'		
	B:	Dare-mo	nani-mo.	
		Nobody	nothing	
		'Nobody, nothing.'		

We show that multiple fragmentary answers form a so-called 'surprising constituent' in the sense of Takano (2002).

4.1. Clause-Mates

Multiple wh-questions allow for wh-phrases from two different clauses in a single sentence to take the same scope. The wh-question in (21) has a wh-phrase in the matrix clause and another wh-phrase in the subordinate clause. While the wh-question in (21) is grammatical, the multiple fragmentary NCI answer is not.

(21)	A:	Dare-ga	[Mary-ga	nani-o	tabeta-to]	itta-no?
		Who-nom	Mary-nom	what-acc	ate-comp	said-Q
		'Who said that Mary ate what?'				
	B:	* Dare-mo	nani-mo.			
		Nobody	nothing			

Fragmentary multiple NCI answers are potentially possible out of a subordinate clause.

(22)	A:	John-ga	[dare-ga	nani-o	tabeta-to]	itta-no?
		John-nom	who-nom	what-acc	ate-comp	said-Q

'Who did John say ate what?'

B: Dare-mo nani-mo.
Nobody nothing

Thus, it is not the fact that one of the wh-phrase is subordinated that makes the answer in (21) impossible. The requirement is that the wh-phrases corresponding to multiple fragmentary NCIs must be clause-mates. The next subsection assimilates this observation to a condition on surprising constituents.

4.2. Surprising Constituents

One distinguishing property of the Japanese language is that it permits multiple phrases in the focus position of a cleft construction. As its literal English translation shows, English does not allow corresponding multiple foci in cleft sentences (Takano (2002)).

(23) John-ga t t watasita-no-wa Mary-ni hana(-o)-da.
-nom passed-NM-top -dat flower-acc-cop
'It was to Mary flowers that John passed t t.'

Takano's explanation is that the multiple foci form a constituent, with one focus phrase adjoined to another.

(24) hana-o
Mary-ni hana-o

We assume this account of (23), and refer to multiple foci forming a constituent as a surprising constituent. (25) shows that constituents from a subordinate clause can form a surprising constituent.

(25) Tom-ga [John-ga t t watasita-to] itta-no-wa Mary-ni hana(-o)-da.
-nom -nom passed-comp said-nm-top -dat flower-acc-cop
'It was to Mary flowers that Tom said that John passed t t.'

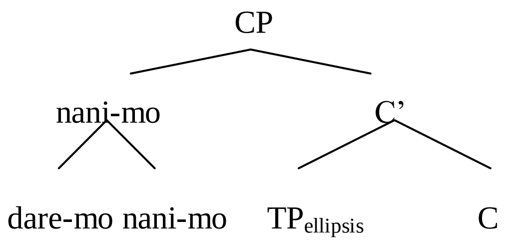
It is not possible, however, for two phrases out of two different clauses to form a constituent.

(26) *t [John-ga Mary-ni t watasita-to] itta-no-wa Tom-ga hana(-o)-da.
-nom -dat passed-comp said-nm-top -nom flower-acc-cop

‘It was Tom flowers that t said that John passed t to Mary.’

We will not go into the details of this account. The point being made here is that surprising constituents, like fragmentary NCIs, observe the clause-mate condition. Based on this observation, we assume that multiple fragmentary NCIs form a surprising constituent. The proposed structure is shown in (27).

(27)



Let us now focus on the order of multiple NCIs in the CP-spec.

4.3. Scrambling and Fragmentary Answers

In a negative answer to a multiple wh-question, the order of the two negative quantifiers must be parallel to the order of the two wh-phrases in the antecedent wh-question. The subject wh-phrase in the wh-question in (28) precedes the object wh-phrase. Switching the order between the subject and the object between the two in the negative answer gives rise to ungrammaticality. This contrasts with (4).

(28) A: Dare-ga nani-o tabeta-no?
who-nom what-acc ate-Q
‘Who ate what?’

B: * Nani-mo dare-mo
Nothing nobody
‘Nothing, nobody.’

Interestingly, when one wh-phrase is scrambled over the other, the negative quantifiers in the negative answer can appear in either order.

(29) A: Nani-o$_i$ dare-ga t$_i$ tabeta-no?
what-acc who-nom ate-Q
‘Who ate what?’

B: Dare-mo nani-mo.
Nobody nothing

B': Nani-mo dare-mo
Nothing nobody

In the multiple wh-question in (28), the subject wh-phrase, *dare-ga*, asymmetrically c-commands the object wh-phrase, *nani-o*. When a phrase A is adjoined to another phrase B, A asymmetrically c-commands B (Kayne (1994)). Thus, in the CP-spec of (27), *dare-mo* c-commands *nani-mo*. This means that the c-command relation between the wh-phrases in the antecedent is preserved between the NCIs in the fragmentary answer. Let us assume that such c-command relations must be preserved (Huang (1982), Richards (2001), Watanabe (1992)). In the fragmentary NCI answer in (28), *nani-mo* asymmetrically c-commands *dare-mo*, but this is not consistent with the c-command relation in the wh-phrases of the antecedent clause, since in the antecedent wh-question in (28), *dare-ga* asymmetrically c-commands *nani-o*.

The wh-question in (29) has scrambled wh-phrases, with the object wh-phrase scrambled over the subject wh-phrase. The scrambled object wh-phrase, *nani-o*, c-commands the subject wh-phrase, *dare-ga*. This c-command relation is consistent with the fragmentary answer B' in (29). The subject wh-phrase in the antecedent, *dare-ga*, also c-commands the trace/copy of the scrambled object. This relation is preserved in the B. Thus, one of the two c-command relations is consistent with B and B'. It is therefore expected that both of the answers are possible in (29). In this way, the observation in (28) and (29) can be accounted for by independent principles of the grammar.

When the wh-phrase is scrambled out of a wh-island, the result is perfectly grammatical, as seen in (30), which is a scrambling counterpart of (8) (Tanaka (1999)). The fragmentary NCI answer is also possible in this example.

(30) A: Nani-o_i John-ga [Mary-ga t_i tabeta-kadooka] siritai-no?
what-acc -nom -nom ate-whether want-to-know-Q
'What does John want to know whether Mary ate?'

B: Nani-mo.
nothing

This is in clear contrast to (8), which does not permit fragmentary NCI answer.

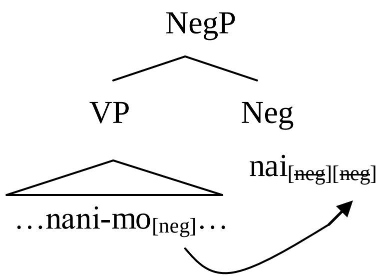
The contrast between (8) and (30), as well as the one between (28) and (29), shows that ellipsis is sensitive to semantically vacuous operations like scrambling, which does not establish an operator variable relation (see Saito (1989), and more recently Saito (2003)). We thus conclude that any accounts of ellipsis based solely on semantics are untenable. Two different accounts of fragmentary NCI answers have been proposed, which both rely on a semantic identity condition on ellipsis proposed in Merchant (2001). We now critically examine their proposals.

5. *Previous Analyses*

5.1. The Feature Copying Account (Watanabe (2004))

Watanabe's account of fragmentary NCI answers assumes that negative quantifiers are inherently negative. Their negative feature is copied to the negative head which itself is negative. The two [neg] features on the Neg head cancel each other out, turning the entire NegP to positive.

(31)



Watanabe is concerned with the condition on ellipsis behind the fragmentary NCI answers in (3). The full-fledged counterpart of (3) is shown below. The phrase in < > is elliptical.

(32) A: Dare-ga pizza-o tabeta-no?
who-nom pizza-acc ate-Q
'Who ate pizza?'

B: Dare-mo < pizza-o tabe-nakatta>.
Nobody pizza-acc eat-neg-past

The problem here is that, while the antecedent wh-question is positive, the presence of *dare-mo* in the answer requires a negation. Ellipsis is possible under some identity condition, but there is no obvious identity between the positive antecedent wh-question and the elliptical negative phrase. Watanabe's idea is that the ellipsis is determined entirely by a semantic identity condition (Merchant (2001)). Merchant's proposal is summarized here.

(33) A constituent α can be deleted only if α is e-given.

(34) An expression E counts as e-given iff E has a salient antecedent A, and module $\exists$-type shifting,
   a. A entails the F-clo(E) and
   b. E entails the F-clo(A)

(35) The F-closure of α, written as F-clo(α) is the result of replacing focus-marked parts with if α with $\exists$-bound variables of he appropriate type (module $\exists$-type shifting).

The F-closure antecedent wh-question in (32) is

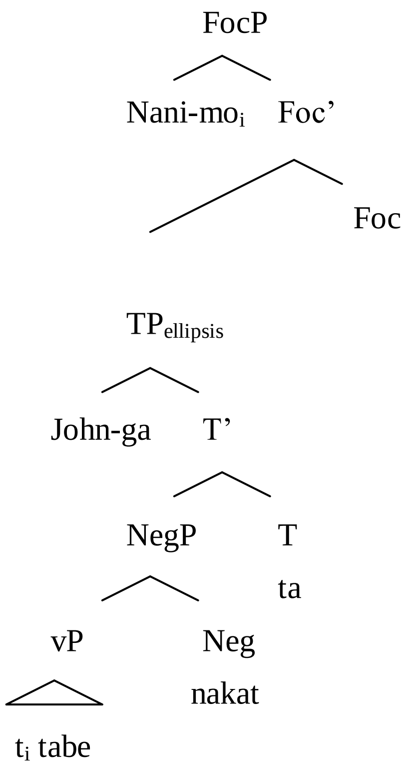
(36) F-clo($\mathrm{TP}_{\mathrm{antecedent}}$) = $\exists$x. x ate pizza

Watanabe (2004:571) remains silent about what category deletes in fragmentary NCI answers. His principal claim is that the two negative features on the Neg head cancel each other out, and turn the entire sentence positive. Since every constituent in the ellipsis site deletes except for the NCI, the category that deletes must be as big as NegP. This would require removing the NCI out of the ellipsis site, in particular, out of the NegP. This is especially important in order to account for fragmentary NCIs in examples like (37), where the subject phrase, which precedes the wh-phrase in the antecedent, deletes along with the rest of the sentence, leaving the fragmentary NCI.

(37)	A:	John-ga	nani-o	tabeta-no?
		John-nom	what-acc	ate-Q
		‘What did John eat?’		
	B:	Nani-mo < John-ga		tabe-nakatta>.
		Nothing	John-nom	eat-neg-past

Hence, the elliptical category must be higher than the projection that dominates the subject, or the past tense, and the NCI must be outside the ellipsis site. In his footnote 13, Watanabe (2004) alludes to a possible analysis along this line in terms of Foc(us)P (Rizzi (1997)). FocP immediately dominates Top(ic)P, which in turn immediately dominates TP, according to Rizzi (1997). Ignoring TopP, the fragmentary NCI answers must have the structure in (38).

(38)



Note that Watanabe's condition on ellipsis cannot require syntactic identity. This is because while the antecedent wh-question in (32) is positive, and therefore lacks NegP, the fragmentary answer in (32) must contain NegP. The structure that contains NegP cannot be syntactically identical to the one that does not. Hence, the hypothesis that ellipsis is conditioned by semantics is crucial. Since the copied [neg] feature cancels the one on the Neg in (31), the F-closures for the antecedent TP and the one for the elliptical TP in (37) are identical.

(39) F-clo($TP_{antecedent}$) = ∃x. John ate x
(40) F-clo($TP_{ellipsis}$) = ∃x. John ate x

One problem with Watanabe's account is that it is not clear how scrambling, a semantically vacuous operation, affects the possibility of fragmentary negative answers. The contrast between (28) and (29), as well as the one between (8) and (30), shows that scrambling remedies potentially ungrammatical fragmentary answers. The antecedent wh-question in (8) is synonymous with the antecedent wh-question in (30), and therefore, the one in (30) could serve as the antecedent for (8), if ellipsis was conditioned purely by semantics.

5.2. Negative Concord as Universal Quantifiers (Giannakidou (2000))

In a series of works, Giannakidous (1998, 2000, 2006) advances the hypothesis that negative quantifiers in Greek and other languages actually do not have a negative feature. Her account is that they are actually universal quantifiers that must scope over the negation. One of Giannakidou's arguments against the negative quantifier analysis is based on (41).

(41) *Dare-mo pizza-o tabeta.
Nobody pizza-acc ate

'Nobody ate pizza.'

Her criticism is that the negative quantifier cannot contribute to the negative meaning of the sentence on its own, as seen by the ungrammaticality of (41), therefore, is not negative. It is clear that this criticism has a problem, since the same criticism applies to any analysis of the negative quantifiers. For instance, consider Giannakidou's own hypothesis that negative quantifiers are actually universal quantifiers, taking scope over negation. If *dare-mo* in (41) was a universal quantifier, (41) should simply mean what (42). The minimal difference between (41) and (42) is that (42) has a Case-marked universal quantifier.

(42) Dare-mo-ga pizza-o tabeta.
everybody pizza-acc ate

'Everybody ate pizza.'

One could claim, following Giannakidou's argument, that *dare-mo* is not a universal quantifier, since it cannot contribute to the universal meaning on its own. The problem that (41) poses is that the example is ungrammatical. What the ungrammaticality of the example shows is that expressions like *dare-mo* are not grammatical on their own under any interpretation, unless they are in the scope of a negation.

5.3. Back to Our Proposal

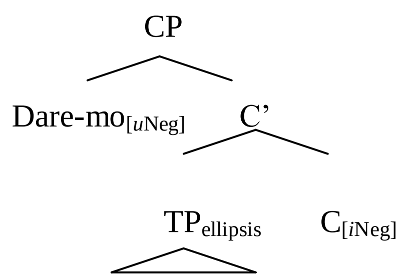
What is necessary in order to account for (41) is a mechanism that rules out *dare-mo* not contained in the scope of a negation. Our account accomplishes this, if *dare-mo* has an uninterpretable feature. In (41), nothing can check the feature. When a c-commanding negation is present, the uninterpretable feature can be checked as long as the neg probes an uninterpretable feature. The uninterpretable feature on the NCI can therefore work as a goal.

(43) *[Dare-mo$_{[u\text{Neg}]}$ pizza-o tabe]-nakatta$_{[i\text{Neg}]}$.
Nobody pizza-acc eat-neg-past

'Nobody ate pizza.'

In fragmentary NCI answers, the NCI moves to the CP-spec. We assume that such NCIs check their uninterpretable neg feature against C, which we assume can be negative.

(44)



In a run-of-mill example of fragmentary NCI answers, the two TPs are identical except for the wh-phrase in the antecedent and the trace/copy bound by the NCI in the ellipsis site.

(45) A: [$_{TPA}$ Dare-ga pizza-o tabeta]-no?
who-nom pizza-acc ate -Q
‘Who ate pizza?’

B: Dare-mo$_i$ <$_{TPellipsis}$ t_i pizza-o tabeta > C$_{[i\text{Neg}]}$
Nobody

In multiple fragmentary NCI answers, both of the NCIs carry an uninterpretable [Neg] feature, with one NCI adjoined to the other. The [*i*Neg] feature on C, being interpretable, does not delete, and therefore, can check the [*u*Neg] feature on both of the NCIs.

(46) A: [$_{TPA}$ Dare-ga nani-o tabeta]-no?
who-nom what-acc ate -Q
‘Who ate what?’

B: [Dare-mo nani-mo] < $_{TPellipsis}$ t t tabeta > C$_{[i\text{Neg}]}$.
Nobody nothing
‘Nobody, nothing.’

Our analysis thus allows us to have a syntactically parallel antecedent for the ellipsis site.

Criticizing Giannakidou’s work, Watanabe (2004:567-8) points out that if a clause that contains a negation can be elided under identity with a positive wh-question, the fragmentary answer in (47) should have the negative interpretation, as specified on the example B. This, however, is not a correct prediction. While the fragmentary answer is possible, it only has the positive interpretation in B’.

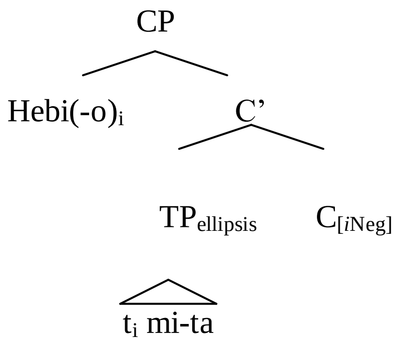
(47) A: Nani-o mita-no?
what-acc see-Q
‘What did you see?’

B: Hebi(-o)$_i$ < *t_i mi-nakatta>.
snake-acc see-neg-past
‘< I didn’t see > a snake.’

B’: Hebi(-o)$_i$ < t_i mita>.
snake-acc see-past
‘< I saw > a snake.’

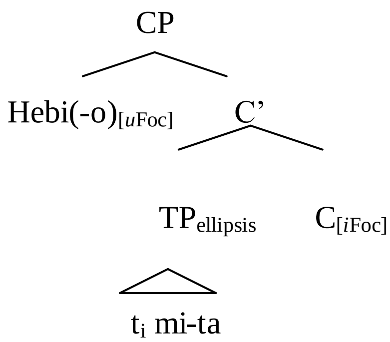
One might think that this criticism also applies to our proposal, if (48) is the structure for the fragmentary answer in (47), and that to this extent, our proposal cannot be maintained, as one of the anonymous reviewers has pointed out.

(48)



However, note that *hebi-o* does not carry a feature that drives movement to the specifier of $C_{[i\text{Neg}]}$, and as such, the structure in (48) does not arise. The proper analysis of (47), fragmentary answers with referential NPs, is beyond the scope of this paper. One possibility is that C can have a focus feature (Miyagawa (2010)), and the referential expression in the fragmentary answer has a focus feature [*u*Foc].

(49)



If, as seems reasonable, $C_{[i\text{Foc}]}$ cannot bear an additional [*i*Neg] feature, we correctly predict the absence of the negative interpretation in (47).

6. Summary

This article has shown that fragmentary negative answers in Japanese involve movement, on a par with the antecedent wh-question. The observation that fragmentary NCIs observe subjacency strongly suggests that movement is involved in the derivation of fragmentary NCIs. Our account is empirically superior to Watanabe’s or Giannadiou’s. It captures the observation that syntactic identity is required for ellipsis. It remains to be seen whether or not the same syntactic condition applies to other elliptical operations in Japanese and other languages.

References

CHOMSKY, NOAM. 1977. "On wh-movement. In *Formal syntax*," eds. PETER CULICOVER, THOMAS WASAW AND AKMAJIAN ADRIAN. New York: Academic press.
CHUNG, SANDRA, LADUSAW, WILLIAM A., & MCCLOSKEY, JAMES. 1995. "Sluicing and logical form." *Natural Language Semantics* 3:239-282.
FIENGO, ROBERT, & MAY, ROBERT. 1994. *Indices and identity*. Cambridge, Mass.: MIT Press.
FOX, DANNY, & LASNIK, HOWARD. 2003. "Successive-cyclic movement and island repair: The Difference between sluicing and VP-ellipsis." *Linguistic Inquiry* 34:143-154.
GIANNAKIDOU, ANASTASIA. 2000. Negative...concord? *Natural language and linguistic theory* 18:457-523.
HUANG, JAMES. 1982. Logical relations in Chinese and the theory of grammar, Department of Linguistics and Philosophy, Massachusetts Institute of Technology.
KAYNE, RICHARD S. 1994. *The antisymmetry of syntax*: Linguistic inquiry monographs ; 25. Cambridge, Mass.: MIT Press.
KISHIMOTO, HIDEKI. 2005. *Toogo koozoo to bunpoo kankei* [*Syntactic structures and grammatical relations*].vol. 8: A series of comparative research on Japanese and English. Tokyo: Kuroshio.
MERCHANT, JASON. 2001. *The syntax of silence: sluicing, islands, and the theory of ellipsis*. Oxford ; New York: Oxford University Press.
MERCHANT, JASON. 2007. "Voice and ellipsis." Ms. Chicago.
MIYAGAWA, SHIGERU. 2010. *Why agree? Why move? Unifying agreement-based and discourse-configurational languages*. Cambridge, Massachusetts: MIT Press.
NISHIGAUCHI, TAISUKE. 1990. *Quantification in the theory of grammar*: Studies in linguistics and philosophy ; v. 37. Dordrecht, The Netherland ; Boston: Kluwer Academic Publishers.
RICHARDS, NORVIN. 2001. *Movement in language*. New York: Oxford University Press.
RIZZI, LUIGI. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. LILIANE HAEGEMAN, 281-337. Dordrecht: Kluwer.
ROMERO, MARIBEL. 1998. *Focus and reconstruction effects in wh-phrases*, Department of Linguistics, University of Massachusetts: PhD dissertation.
ROSS, JOHN ROBERT. 1967. *Constraints on variables in syntax*, Department of linguistics and philosophy, Massachusetts Institute of Technology: PhD dissertation.
ROSS, JOHN ROBERT. 1969. "Guess who?" In *Proceedings of the Fifth Regional Meeting of the Chicago Linguistic Society*, ed. R. BINNICK et al eds, 252-286. Chicago, Illinois.
ROSS, JOHN ROBERT. 1983. "Inner islands." Ms. Cambridge, MIT.
SAG, IVAN. 1976. Deletion and logical form, Department of linguistics and philosophy, Massachusetts Institutes of Technolog: PhD dissertation.
SAITO, MAMORU. 1989. "Scrambling as semantically vacuous A'-movement." In *Alternative conceptions of phrase structure*, eds. MARK BALTIN & ANTHONY KROCH, 182-200. Chicago: The University of Chicago Press.
SAITO, MAMORU. 2003. "A derivational approach to the interpretation of scrambling chains." *Lingua* 113:481-518.
SAUERLAND, ULI. 1996. "Guess how." Paper presented at *Proceedings of CONSOLE*, Leiden, The Netherlands.

TAKANO, YUJI. 2002. "Surprising constituents." *Journal of East Asian Linguistics* 11:243-301.

TANAKA, HIDEKAZU. 1999. "LF wh-islands and the minimal scope principle." *Natural language and linguistic theory* 17:371-402d.

WATANABE, AKIRA. 1992. "Subjacency and s-structure movement of wh-in-situ." *Journal of East Asian Linguistics* 1:255-291.

WATANABE, AKIRA. 2004. "The genesis of negative concord: syntax and morphology of negative doubling." *Linguistic Inquiry* 35:559-612.

Hidekazu Tanaka
Department of Language and Linguistic Science
University of York
Heslington
York
email: hidekazu.tanaka@york.ac.uk