Es gibt eine noch stärkere Form eines Zusammenwirkens: das Ü mit dem J, das ergibt das schwedische Y. Da klingt der konzentrierteste Laut Y mit dem U zusammen. Sie treffen einander weit vorn auf den Lippen, ja man kann sagen, dieser Treffpunkt liegt sogar vor dem Lautorganismus im Übersinnlichen.

Dies ergibt allerdings etwas Besonderes: Es ist die Prim = U und die Septime = I, die da zusammenklingen und eigentlich entsteht da ein Triller! Da spielt Übersinnliches stark mit hinein. Wenn der Mensch dies singt, wird sein Kopf ganz flach, es leuchtet im Ätherischen auf. Aber wenn man da hinkommt, ist man eigentlich schon außer sich.

So könnte man von den Umlauten sagen, dass sie komödienhafte Vokale sind, weil sie fortwährend zwischen zwei Gegensätzen wechseln. Das wirkt komisch. Rudolf Steiner sagte daher: Die Umlaute Ö, Ä, Ü gehören zur Komödie

So ist O ein dicker Laut, E aber ein magerer. Das fortwährende Wechseln zwischen beiden gibt einen Sprung: Ö; das Wort Öhrchen z.B. hört sich doch wie eine Komödie an; dick und dünn zusammen ist doch ein Witz, eine Groteske.

Dagegen können die Diphthonge schon ernst sein, sie geben mehr Bewegung: AU, EI, ÄU.

So könnte man in Bezug auf ihr Wirken noch manches über die ertönenden Vokale sagen.

So ist U der stehende (vor Furcht starr gewordene Mensch). Das I ist der gehende Mensch (Streckung der Beine).

Der U-Mensch ist still, ruhig, aber wenn das I in ihn hineinschlägt, bewegt es den U-Menschen. Der U-Mensch hat Symmetrie in den Beinen, der I-Mensch strahlt! Stehen und Gehen zusammen ist das Ü. Wenn wir E sagen, wird die Wirbelsäule gestreckt, bei I erst recht.

Den Sprachmenschen (Kehlkopf) muss man sich so vorstellen; liegend mit dem Bauch nach unten, mit beiden Armen vorgestreckt, mit den Beinen schwimmend (also seine Beine in unsere Richtung nach vorne). Unsere Kiefer sind die Beinknochen und unsere Lippen sind das Fleisch auf den Beinen (als Wiederholung des physiologisch dreigegliederten Menschen angesehen).

In einem E kommt etwas Gedankenartiges ins Vokalische hinein. Es klingt leicht blechern und ist wie von der Blässe des Gedankens angekränkelt. Das E weist verschiedene metallische Qualitäten auf, aber diese können durch den Klang sehr veredelt werden. Das E kann einen goldähnlichen Klang bekommen. (Dr. Eugen Kolisko nannte das E einen `negativen Verstand')

Das A ist der Anfang der Musik und mit dem L (=Zunge) davor; in LA (das ist die Sexte in der Reihe des Do, Re, Mi) wirkt etwas Urmusikalisches.

Fassen wir nun noch einmal zusammen, so ergibt sich, dass wir eine vierfache Verbindung mit den Vokalen haben: